國 立 交 通 大 學

生物科技研究所

碩 士 論 文

融合一段 SARS 片段至腫瘤來源的胜肽可提高免疫反應

The Enhancement of Immunogenicity by a SARS Fragment Fusion to a Tumor-Derived Peptide

研 究 生：高詩涵

指導教授：廖光文博士

中 華 民 國 九 十 五 年 七 月

融合一段 SARS 片段至腫瘤來源的胜肽可提高免疫反應

The Enhancement of Immunogenicity by a SARS Fragment Fusion to a Tumor-Derived Peptide

研 究 生：高詩涵　　Student：Shih-Han Kao

指導教授：廖光文　　Advisor：Kuang-Wen Liao

國 立 交 通 大 學

生 物 科 技 系

碩 士 論 文

A Thesis

Submitted to Institute of Biological Science and Technology

College of Biological Science and Technology

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Biological Science and Technology

July 2006

Hsinchu, Taiwan, Republic of China

融合一段 SARS 片段至腫瘤來源的胜肽可提高免疫反應

研究生：高詩涵　　指導教授：廖光文　博士

國立交通大學生物科技研究所

中文摘要

腫瘤細胞會表現低免疫反應之腫瘤相關抗原 (TAA)。一段癌胚抗原 (CEA) 在CT26腫瘤模式中被用來模擬低免疫反應之腫瘤相關抗原。我們藉由SARS冠狀病毒的片段修飾之胜肽疫苗提高在Balb/c老鼠動物模式裡的CEA專一療效。並且藉由提高SARS片段上抗原決定位 (epitope) 與主要組織適應性複合體第一型 (MHC class I) 之結合力來確保SARS片段之免疫反應效果。利用網路軟體 (http://www.syfpeithi.de/) 分析顯示，突變越多的抗原決定位表示其與主要組織適應性的結合力越好。這些質體設計分別轉殖入沙門氏桿菌 (*Salmonella typhimurium*) 藉以免疫Balb/c老鼠。細胞生理細胞激素測試(in vitro cytokine assay)顯示只帶有CEA的組別只引起介白質IL-4 的分泌，而其他多帶有SARS片段的組別，不論有無突變點，都能有效的引起腫瘤壞死因子-α (TNF-α) 及介白質IL-10的分泌。活體細胞激素測試 (in vivo cytokine assay) 則顯示單獨的CEA不能引起Th1 與Th2的反應，但多加的SARS片段則可。老鼠在經過免疫後接種腫瘤大小上的表現也以有SARS片段者來得小，甚者其存活率與無腫瘤率也以SARS組較佳，顯示SARS片段，不論有無突變，對CT26 有較好的抑制效果。在治療上面，CEA表現無法有效的抑制腫瘤生長，但SARS則可。

綜合以上結果，低免疫反應之腫瘤相關抗原，如 CEA，是無法在動物模式上面帶來有效的保護效果，而在我們建立的系統裡，一段普遍的抗原—SARS 片段—則可以加強原低免疫反應的腫瘤相關抗原的反應，進而提高 DNA 疫苗的效果。此外，我們也建立了一個平台，可以利用電腦預測的方式提高佐藥疫苗的療效。

The Enhancement of Immunogenicity by a SARS Fragment Fusion to a Tumor-Derived Peptide

Student: Shih-Han Kao Advisor: Dr. Kuang-Wen Liao

Institute of Biological Science and Technology
National Chiao Tung University

ABSTRACT

Tumor cells express tumor-associated antigens (TAAs) that are usually in low immunogenicity. A fragment of carcinoembryonic antigen (CEA) was utilized to simulate the low immunogenous TAA on colon carcinoma, CT26. To enhance the efficiency of our DNA vaccine, it was fused with an exogenous SARS-CoV fragment which is high immunogenous and is expected to induce and enhance the immune response in our animal model, Balb/c mice. The SARS fragment was mutated based on the affinity prediction between the epitope and MHC molecules (H2-Kd) by Internet software (http://www.syfpeithi.de/). The more mutations there are in a construct, the higher the affinity is. These constructs were then transformed into *Salmonella typhimurium* and orally fed to immunize Balb/c mice. In vitro cytokine profile reveals that CEA alone induces only IL-4 secretion whereas constructs with an additional SARS fragments, whether mutated or not, can significantly induce TNF-αand IL-10 secretion. In vivo cytokine profile shows that CEA alone can not induce any cytokine secretion but an additional SARS fragment fused to CEA can induce both Th1 and Th2 responses. Mice in the protection assay also had smaller tumor volume than those with CEA alone. The efficiency of the CEA-SARS immunization (both non-mutated and mutated) is reflected on their survival rate and tumor-free rate, which are both higher than the CEA alone group. Moreover, the SARS fragment fused with CEA effectively slowed tumor growth in the therapeutic assay. In conclusion, low immunogenous TAAs, such as CEA, can not effectively induce the immune response of animal models. We have set up a system in which the foreign parental or mutated SARS fragments could enhance the anti-tumor efficacy of the tumor vaccine against endogenous tumor antigens. Furthermore, we provide a platform to enhance the adjuvant effect of the foreign peptide by computer prediction.

Acknowledgement

我要感謝許多人，因爲有了他們才使得我的論文能夠完成。首先要感謝的是我們實驗室的大家長—廖光文老師，兩年前他不計我的背景願意收我當他的指導學生，也信任我可以在這兩年內可以完成實驗和修課內容，因爲有了這樣的契機使我得以實現自己的理想，而每當我實驗遇到困難時，他總是熱心與我討論，謝謝他的指導讓我從什麼都懵懂無知到開始欣賞科學的嚴謹和邏輯。再來要感謝的是靜宜學姐，總是在實驗實務操作上給我建議和指導，並且帶著我做了許多實驗；另外還要感謝同組的學弟：侑松和昱丞，每每爲了毒殺實驗必須陪著我早起；源庭要分擔流式細胞儀的分析，其翰很辛苦地打了無數盤的 ELISA，真的很謝謝你們！

另外還要感謝我的同學上知，他陪我去了很多趟鼠房也陪我做了很多實驗，還幫我們買了便當，以在大家實驗做得焦頭爛額的同時得以溫飽；還要感謝竣元和弘育，在午夜的時分幫我收膠或收菌。感謝懷堯、韻如、彥谷等學弟妹日常生活中的照顧，爲枯燥不變的生活增添幾分色彩！還要感謝我的實驗老鼠，每次看到他們被犧牲都覺得很不捨，我希望每一條生命的犧牲都有其代價。

最後還要感謝我的父母，因爲有家人的支持和鼓勵，我才能順利地完成學業。感謝所有曾經幫助過我的朋友，因爲有你們的相左，讓我瞭解友誼的真諦。

詩涵 於 2006

Contents

Tables

Figures

Abbreviations

APCs	antigen-presenting cells
CEA	carcinoembryonic antigen
CTLs	Cytotoxic T lymphocytes
DCs	dendritic cells
DIOC18	3,3'-dioctadecyloxacarbocyanine
DMEM	Dulbecoo′s modified eagle medium
Th cells	Helper T cells
IFN-γ	interferon-γ
IL-10	interleukin-10
IL-12	interleukin-12
IL-2	interleukin-2
IL-4	interleukin-4
IL-5	interleukin-5
kDa	killoDalton
MHC	major histocompatibility complex
NK cells	natural killer cells
PBS	phosphate buffer saline
PI	propidium iodide
SARS-CoV	severe acute respiratory syndrom coronavirus
TAA	tumor-associated antigen
TCR	T-cell receptor

Chapter 1 Introduction

1.1 Tumor antigens

The immune system surveys the body for abnormal antigens or cells not only of infectious microorganisms but also of tumor. It has been observed that tumor expresses tumor associated antigens (TAAs) that can be recognized and serve as the target for immune cells such as cytotoxic T cells (Boon, Cerottini et al. 1994; Pardoll 1994). Six categories of TAAs can be defined due to the different mechanisms that result in the generation of TAAs and described in the following sections.

1.1.1 TAAs from oncogene mutations or rearrangement

These antigens are the result of point mutations or gene rearrangements, which often arise as part of the process of oncogenesis. For example, the transforming allele of the Ki-ras2 gene present in the human colon carcinoma cell line SW480 differs from its normal counterpart only at the amino acid at position 12. The normal cDNA encodes a glycine at this position, and the transforming allele encodes a valine. Expression of these cDNAs indicates that this amino acid 12 alteration confers oncogenic activity on the mutated gene (McCoy, Bargmann et al. 1984). Chronic myelogenous leukemia (CML) is a human disease associated with a consistent chromosomal translocation that results in sequences from the c-abl locus on chromosome 9 being fused to sequences in a breakpoint cluster region (bcr) on chromosome 22 (Ben-Neriah, Daley et al. 1986).

1.1.2 TAAs from mutated tumor-suppressor gene products

Normal cells contain repressors whose loss result in uncontrolled growth (King 2000). The p53 gene, for instance, which produces the p53 repressor protein, is the gene most

frequently altered in human cancers (King 2000). The p53 gene is frequently mutated or inactivated in all types of human lung cancer. The genetic abnormalities of p53 include gross changes such as homozygous deletions and abnormally sized messenger RNAs along with a variety of point or small mutations, which map to the p53 open reading frame and change the amino acid sequence in a region highly conserved between mouse and human (Takahashi, Nau et al. 1989).

1.1.3 TAAs from reactivated embryonic gene products not expressed in adult tissues

Some TAAs, which are the embryonic gene products, are somehow turned on in tumors. For instance, MAGE-1 and MAGE-3 are two clinically relevant antigens expressed in many human melanomas and other tumors, but not in normal tissues, except testis (Bueler and Mulligan 1996). Alpha-fetoprotein is another example whose mRNA levels increase in hepatocellular carcinoma (HCC) cells as compared with non-neoplastic tissues (Matsumura, Ijichi et al. 2001).

1.1.4 TAAs from viral gene products

Some viruses can provide an oncogene which codes for a functional product, as in the case of Rous sarcoma virus. The oncogene v-src codes for a 60 kDa phosphoprotein ($pp60^{src}$) that has tyrosine kinase activity (King 2000). Other viruses may insert a part of the sequence into a host's genome. For example, the regulatory sequences repeated at each end of the viral genome of mouse mammary tumor virus (MMTV) enhance the transcription of the nearby genes (King 2000). Still others can act directly by virtue of their v-onc protein product binding to and inactivating host proteins, such as human papilloma virus. Inactivation of tumor suppressors, Rb and p53, by viral E7 and E6 proteins is required for cervical cancer (King 2000).

1.1.5 TAAs from idiotypic epitopes

Some of the TAAs are specifically expressed on tumors but not on normal cells, including idiotypic epitopes. One of the best characterized idiotypic epitopes is the idiotypic immunoglobulin (Id) of B-cell lymphoma. The Id is determined by the rearrangements of the variable heavy (VH) and light (VL) chains of the immunoglobulin V regions that are unique for each clonal B-cell population and represent tumor-specific antigens (Muraro, Bondanza et al. 2005).

1.1.6 TAAs from tissue-specific self antigens expressed by tumors

The tissue-specific self antigens are specifically expressed on tumors but not on normal cells as well. They can be exemplified by carcinoembryonic antigen (CEA), which is a membrane-anchored glycoprotein expressed on the great majority of colorectal, gastric, and pancreatic carcinomas as well as 50% of breast cancers and 70% of non-small cell lung cancers (Thompson, Grunert et al. 1991). Another example of self antigens is the tyrosinase of melanoma, which is responsible for the synthesis of melanin in the epidermis via oxidation of tyrosine (Halaban, Pomerantz et al. 1983). Melanin biosynthesis (melanogenesis) is a metabolic pathway exclusively expressed by melanocytes and melanoma cells, and is often altered and/or markedly elevated in the latter cells (Salopek and Jimbow 1996). Therefore, tyrosinase is overexpressed by melanoma and serves as a self antigen for clinical detection (Kounalakis and Goydos 2005).

1.2 MHC types and antigen-presenting pathways

Tumor antigens, whether endogenous or exogenous, are all presented by the major histocompatibility complex (MHC) molecules to different T cells. There are mainly two types

of MHC molecules, MHC class I and MHC class II, each of which functions in different ways.

1.2.1 MHC class I molecules and their functions

MHC class I molecules, expressed on the surface of the all cell types except for red blood cells (RBCs), present peptides derived from endogenous antigens to CD8+ cytolytic T lymphocytes (CTL). Degraded by proteasomes, these peptides bind to MHC class I molecules in the lumen of the endoplasmic reticulum via the transporters associated with antigen processing-1 and -2 (TAP1 and TAP2). The presentation of peptides by the MHC class I molecules initiate the activation of CTL, which is one of the effector cells in antitumor immunity.

1.2.2 MHC class II molecules and their functions

MHC class II molecules, expressed on the antigen-presenting cells (APCs), such as B cells, dendritic cells, and macrophages, present peptides derived from exogenous antigens to CD4+ helper T (Th) cells. These tumor antigens are released from tumor cells, taken up and processed into peptides by APCs, then displayed on MHC class II molecules. The recognition by Th cells releases cytokines, such as IL-2, to promote the activation of CTLs (Li, Zhang et al. 2006).

1.3 Binding affinity and presentation strength

When molecules interact, binding affinity accounts for the strength between them. In the case of the MHC molecule presentation to T cells, there are two kinds of binding affinity that need to be considered to activate the T cell response, namely the binding affinity between the epitope and the MHC molecule and that between the MHC-epitope complex and the T-cell

receptor (TCR) of a T cell.

1.3.1 MHC Class I binding of its peptides

Because tumor antigens are mostly recognized by the CD8+ T cells, it is of importance to introduce the binding between the peptide and the MHC class I molecule.

The binding of a peptide in the peptide-binding cleft of an MHC class I molecule is stabilized at both ends by contacts between atoms in the specific residues of the peptide and invariant sites that are found at each end of the cleft of all MHC class I molecules through a series of hydrogen bonds and ionic interactions. Peptides that bind to MHC class I molecules are usually 8-10 amino acids long. The peptides that can bind to a given MHC variant have the same or very similar amino acid residues at two or three particular positions along the peptide sequence. The side chains of the amino acids at these positions insert into pockets in the MHC molecule that are lined by the polymorphic amino acids. Because the binding of these side chains anchors the peptide to the MHC molecule, the peptide residues involved have been called anchor residues, which are usually hydrophobic at the carboxyl terminus. In some cases, particular amino acids are preferred in certain positions, whereas in others the presence of particular amino acids prevents binding (Janeway 2005).

1.3.2 MHC Class I binding affinity and presentation strength

Although some tumor antigens can elicit the T-cell response, most of the endogenous T cell responses are weak. This is because the affinity between the peptide antigen and its presenting MHC molecule is generally weak (Cox, Skipper et al. 1994). Such antigens are often characterized by an absence of favored residues at critical anchor positions involved in MHC binding. For weak tumor antigens that fall into the low-affinity MHC binding category, replacement or mutation of unfavorable anchor residues with more effective ones may greatly enhance MHC binding properties (Lurquin, Van Pel et al. 1989; Gervois, Guilloux et al. 1996;

Parkhurst, Salgaller et al. 1996; Bakker, van der Burg et al. 1997). These altered peptides may more effectively activate T cell responses against the wild-type peptide antigen by virtue of the increased efficiency of presentation of the MHC-peptide complex to specific T cells. In both mouse and human studies, these anchor-modified peptides can elicit superior T cell responses against the original antigen in vivo (Dyall, Bowne et al. 1998; Rosenberg, Yang et al. 1998).

1.3.3 The binding affinity between the MHC-epitope complex and the T cell

Some antigens bind to their presenting MHC with affinities in a similar range to the viral antigens yet elicit weak endogenous immune responses (Lee, Yee et al. 1999). For tumor antigens with high MHC affinities, the proposed mechanism for weak endogenous immune responses is that high-affinity T cells are actively tolerated via anergy or deletion, thereby leaving a functional repertoire consisting of T cells bearing T cell receptors (TCR) with low affinity for MHC-peptide complexes. This residual T cell repertoire is postulated to have escaped active tolerance induction by virtue of its low affinity for MHC-peptide ligand. This mechanism is particularly relevant for shared tumor antigens, which, because they are self-antigens, have had a long period of time to induce tolerance (Morgan, Kreuwel et al. 1998).

The low binding affinity between the MHC-epitope complex and the TCR can be elevated by amino acid substitution. Slansky et al indicated in their research that the improved immunity results from enhanced in vivo expansion of T cells specific for natural tumor epitopes (Slansky, Rattis et al. 2000).

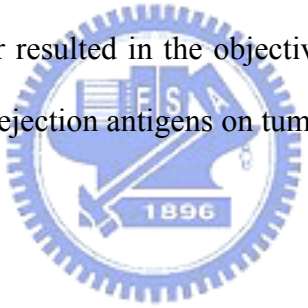
1.4 Tumor vaccines

TAAs have been utilized to induce antitumor immunity in tumor vaccines. Several

strategies have been developed to utilize TAAs for tumor vaccines (Rosenberg 1996; Chen and Wu 1998; Timmerman and Levy 1999). Some of the most important tumor vaccines will be discussed below.

1.4.1 Tumor-infiltrating lymphocytes (TILs)

The earliest discovery of tumor-infiltrating lymphocytes has suggested the involvement of immune cells in the suppression of tumor growth. Studies in animal tumor models have indicated that progressively growing tumors contain weak or nonreactive tumor lymphocytes but that regressing tumors have highly reactive lymphocytes (Gillespie and Russell 1978). In the study by Wang and Rosenberg, the adoptive transfer of cytotoxic T lymphocytes (CTLs) derived from tumor-infiltrating lymphocytes (TIL) along with interleukin-2 (IL-2) into autologous patients with cancer resulted in the objective regression of tumor, indicating that these CTLs recognized cancer rejection antigens on tumor cells (Wang and Rosenberg 1996).



1.4.2 Tumor antigen fusion

Fused tumor antigens have been constructed to elicit antitumor effect. A recent research has indicated that mucin 1 (MUC1), a TAA for epithelial carcinoma, when fused with heat shock protein complexes (HSP65) isolated from tumors, can induce specific and nonspecific antitumor immunity. HSP65-MUC1 induces growth inhibition of MUC1-expressing tumors and increases survival of the tumor-bearing mice by activating MUC1-specific CTL and enhancing IFN-γ secretion (Li, Li et al. 2006).

1.4.3 Cytokine fusion

Other researchers enhance the immunity by cytokines that have been well-known for their proinflammatory or T-cell mediated responses. For example, the human IL-2 has been fused with human MUC1, which is over-expressed and aberrantly glycosylated in most breast

tumors. The patients vaccinated intramuscularly with a single dose of the recombinant vaccinia virus resulted in no significant clinical adverse effects. None of the nine patients had a significant increase in MUC1-specific antibody titers after one single injection (Scholl, Balloul et al. 2000).

1.4.4 Antigen-presenting cells (APCs) fusion

APCs can prime naïve T cells and initiate a prime immune response (Steinman 1991; Banchereau and Steinman 1998). Among all, dendritic cells (DCs) are one of the most utilized APCs to induce a T-cell response. Various DC-based strategies, such as DCs pulsed with tumor-associated peptides or proteins, viral transduction of DCs with tumor-specific genes or transfection with liposomal DNA or RNA, have been developed to introduce tumor specific antigens into DCs and thereby to generate cytotoxic T lymphocyte (CTL) responses against malignant cells (Boczkowski, Nair et al. 1996; Condon, Watkins et al. 1996; Gong, Chen et al. 1997; Ribas, Butterfield et al. 1997; Song, Lee et al. 1997; Song, Kong et al. 1997; Specht, Wang et al. 1997). Xu et al. has generated the fusion between human cancer SW480 cells and DCs to elicit interferon-γ secretion against colon cancer (Xu, Ye et al. 2004).

1.5 Vaccine approaches

Approaches to deliver tumor vaccines can vary, but they can be categorized into two main groups by their molecular components, protein injection and DNA vaccine. The previous group is later subcategorized into two subgroups. One utilizes dead tumor cells as a vaccine approach, and the other utilizes TAAs or any other protective molecules in the form of functional proteins.

1.5.1 Irradiated tumor cells as a vaccine approach

Irradiated tumor cells were once injected into animal models (Mazurek and Duplan 1959; Shibata, Jerry et al. 1976). However, such an approach may only exert a limited effect. Therefore, they have been engineered later with other molecules to exert a synergistic effect on tumors. For example, in a more recent study, Jain et al. transduced irradiated CT26 cells with granulocyte-macrophage colony-stimulating factor to enhance the efficacy of their tumor vaccine. Along with the help of systemic injection of IL-2, 88% of the treated mice were tumor free on day 21 (Jain, Slansky et al. 2003). Besides, irradiated tumor cells can be incubated with DCs and uptaken by them as a vaccine approach. Such pulsed DCs have been proved to elicit the immune response successfully. The specific T cell responses were observed. In vitro studies showed that fusions effectively activated CD8+ T lymphocytes to secrete IFN-γ (Xu, Ye et al. 2004).

1.5.2 Protein vaccine

Vaccines can be in the form of proteins. Cytokines, such as IL-2, have been injected intratumorally with TILs in a breast cancer model. In the study of Liu DL et al., such an immunotherapy with rIL-2 and TILs were given to Wistar rats with breast cancer. The total response rate was 42%, of which 25% tumors showed partial regression and 17% tumors reached complete remission where infiltration of plenty of T lymphocytes was detected, indicating that T cell-mediated antitumor immunity is primarily responsible for tumor rejection. (Liu, Yang et al. 1996).

Pulsed APCs can be injected into animals to elicit immune responses against tumors. For instance, DCs can be pulsed with an antigenic peptide or tumor lysate to induce specific CTLs against tumors. In the study of Waeckerle-Men et al, autologous DCs were pulsed with multiple T cell epitopes derived from four different prostate-specific antigens in patients with advanced hormone-refractory prostate cancer. The vaccination elicited significant cytotoxic T cell responses against all prostate-specific antigens tested. In addition, memory T cell

responses against the control peptides derived from influenza matrix protein and tetanus toxoid were efficiently boosted (Waeckerle-Men, Uetz-von Allmen et al. 2006).

1.5.3 DNA vaccine

DNA vaccine, which is to construct the desired DNA segment in a vector to elicit immunity, is another approach for vaccine delivery. The desired DNA segment can be a TAA, cytokine, or the combination of both in the form of DNA sequence. For instance, *neu* is an oncogene for breast cancer. Chen et al generated DNA expression vectors encoding the full-length *neu* cDNA (designated pNeuN), the neu extracellular domain (pNeuE), and the *neu* extracellular and transmembrane domains (pNeuTM). The intramuscular injection of pNeuTM or pNeuE, and to a lesser extent pNeuN, induced protective immunity against a subsequent challenge with Tgl-1 cells, a neu-expressing tumor cell line, in FVB/N mice (Chen, Hu et al. 1998).

Cytokine can be delivered into animals in the form of DNA instead of protein. In the study of Schultz J et al., the intramuscular injection of plasmid DNA coding for IL-12 abolished the establishment of pulmonary metastases of B16F10 melanoma cells in a syngeneic mouse model. Moreover, it also resulted in a pronounced reduction of tumor growth in C57/BL6 mice. This antitumor effect correlated with a long-lasting expression of cytokines, which manifested itself as high levels of IL-12 in the serum 12 days after DNA treatment (Schultz, Pavlovic et al. 2000).

However, TAA alone or cytokines cannot always induce enough immunity against tumor. Combined therapy has been adopted in some of the therapeutic approaches. Marshall et al. tested a prostate-specific antigen (PSA) DNA vaccine along with the co-administration of pIL-18 plasmid in a mouse tumor model. Complete tumor protection mediated by both CD4+ and CD8+ T cells was observed in all mice. Analysis of the immune response in mice immunized with either pPSA or pPSA/pIL-18 demonstrated that pIL-18 skewed the

PSA-specific immune response toward Th1. More importantly, stronger CD4+ and CD8+ T cell responses developed in the pPSA/pIL-18-immunized mice with faster kinetics (Marshall, Rudnick et al. 2006).

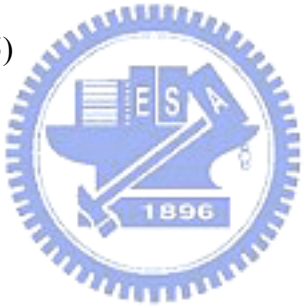
1.6 Oral DNA vaccine

DNA vaccine, though usually injected intramuscularly, can be administered into the animal model orally by bacteria. It has been observed that some of them display preferential replication or preferential accumulation in the tumor microenvironment (Sznol, Lin et al. 2000). In contrast to viruses, the bacteria reside primarily in the extracellular tumor microenvironment and possess certain features that may be advantageous in the treatment of cancer (Kops 1997). Moreover, because of their large genome size, bacteria can readily express multiple therapeutic transgenes, such as cytokines or pro–drug-converting enzymes, and their spread can be controlled with antibiotics if necessary (Sznol, Lin et al. 2000).

Among all, *Salmonella typhimurium* has been one of the most utilized bacteria to carry therapeutic transgenes. In 1997, Pawelek et al. reported that *Salmonella* would infect and preferentially accumulate within implanted tumors in mice, achieving tumor/normal tissue ratios of approximately 1,000:1 (Pawelek, Low et al. 1997). However, wild-type *Salmonella* contains certain virulence and toxicity that when administered systemically, it may threaten the safety of the host (Clairmont, Lee et al. 2000). Therefore, to develop a clinical candidate with a high safety profile, a wild-type strain of *Salmonella typhimurium* was attenuated by partial deletion of the *msbB* gene, whose product is responsible for addition of a terminal myristic acid group in the formation of lipid A (Somerville, Cassiano et al. 1999). Lipopolysaccharide derived from these lipid A mutants is markedly diminished in ability to induce TNF- ˙ in vitro in human monocytes and in vivo after administration to mice and pigs (Low, Ittensohn et al. 1999). As an additional safeguard, deletion of the *purI* gene

(requirement for an external source of purine, e.g., adenine) was engineered into the *msbB*–*Salmonella* strain as a second attenuating mutation (Low K 1999; Low, Ittensohn et al. 1999; Luo X 1999). The gene modifications do not affect the ability to achieve high tumor/normal tissue ratios in mouse models, and the bacteria maintained their capacity to inhibit the growth of both subcutaneous tumors and lung metastatic diseases (Sznol, Lin et al. 2000).

Other modifications have been reported. For example, the strain used in this study is the attenuated aromatic acid-dependent (aro) *Salmonella typhimurium*, which has been well characterized as carriers for various heterogeneous antigens (Dougan G. 1986; Fagan, Djordjevic et al. 1997). These vaccine strains are capable of colonizing the gut-associated lymphoid tissues (Peyer's patches) and secondary lymphatic tissues including spleen and liver following oral administration in mice, to elicit mucosal, humoral and cell-mediated immune responses (Hormaeche CE 1995)



1.7 Strategy

1.7.1 The concept of design

Tumor-associated antigens have some disadvantages. First of all, they are usually in low immunogenicity due to their low binding to the MHC class I molecule (Cox, Skipper et al. 1994). Second, they can be toxic or tumorigenic themselves if not properly handled. For example, the E6 and E7 proteins of the HPV bind p53 and Rb, respectively and inactivate them (King 2000). As a result, a full-length TAA is not a practical approach for a therapeutic target. In contrast, a peptide, if used in the vaccine, can reduce the toxicity or tumorigenicity of the TAA. Yet, peptides are supposedly to be less immunogenous because they generate fewer epitopes than an antigen after proteasome degradation. Therefore, an exogenous peptide that is immunogenous is considered in our strategy to enhance the immunity of our animal model.

Viruses are a suitable source of exogenous peptides. Many viruses can elicit strong immunity in humans. For example, severe acute respiratory syndrome coronavirus (SARS coronavirus) has been proved to elicit a cytokine storm, in which IFN-γ, IL-18, TGF-β, IL-6, etc., are elevated significantly (Wong, Lam et al. 2004; Zhang, Li et al. 2004; Huang, Su et al. 2005). However, the immune response should not exceed the threshold, beyond which viral infection will cause great damage to the host. To solve this problem, a fully replicable virus can not be utilized in the experiment.

1.7.2 Strategy

Here, the construction of a fragment of the SARS coronavirus (SARS-CoV) was preceded by a low immunogenous peptide, carcinoembryonic antigen (CEA) on a vector. The selected SARS fragment has been verified to effectively elicit IFN-γ secretion in human (Xu, Ye et al. 2004). The most possible epitope for Balb/c mice within this SARS peptide was calculated by the SYFPEITHI (http://www.syfpeithi.de/), which is an affinity-predicting website between the MHC molecule and the epitope presented. The higher the score is, the better the affinity predicted by the computer is.

The epitope that has the highest score is modified by point mutation to yield an epitope with even a higher score (m1). Subsequently, the amino acids of this m1 epitope is further mutated by SYFPEITHI to obtain a much higher score sequence (m2). Finally, a third epitope that has another cumulative point mutation is calculated by the SYFPEITHI (m3), which has the highest score among all.

Each plasmid was designed to transform *Salmonella typhimurium* to generate an immune response within tumor. The genetically engineered *Salmonella* may stimulate and activate naïve T cells. After immunization with *Salmonella* transformed with pAAV-CEA, pAAV-CEA-SARS, pAAV-CEA-m1, pAAV-CEA-m2, and pAAV-CEA-m3, respectively,

CT26 cells producing the peptide CEA (CT26/CEA) were inoculated into mice and tumor sizes were measured. In addition, the immune responses, such as cytotoxic activities and cytokine releases of splenocytes, were analyzed. These results showed that the peptide derived from the SARS virus and fused with the low immunogenous tumor antigen can strengthen the immune activities of the host against tumor cells with CEA peptide production.

Chapter 2 Materials and Methods

2.1 Materials

2.1.1 Primers

Gene	Primer	Sequence (5'→3')	Tm (℃)
CEA	**P1**	TAC GGA ATT CAT GGA GTC TCC CTC GGC CCC TCC CCA CAG ATG GTG CAT CC	71.6
	P2	CCT GGC AGA GGC TCC TGC TCA CAG GTG AAG GGA GGA CAA C	71.5
	P3	CTG GGA GAG GGT GGG AGG AGG GAG CTG GGG TCT CCT GGG T	75.1
	P4	CTC CTC CCA CCC TCT CCC AGG TTG TCC TCC CTT CAC CTG T	71.8
	P5	GAG CAG GAG CCT CTG CCA GGG GAT GCA CCA TCT GTG GGG A	73.7
	P6	GCT ATC TAG ATC ACA GCC CTG TCC TAC CCA GGA GAC CCC AGC TCC	71.1
	P7	GCT ATC TAG ACA GCC CTG TCC TAC CCA GGA GAC CCC AGC TCC	71.0
SARS	**P1**	TAC GTC TAG AAA AGT CGA GGC GGA GGT ACA AAT TGA CAG GTT AAT TAC A	65.7
	P2	GGC AGA CTT CAA AGC CTT CAA ACC TAT GTA ACA CAA CAA C	63.9
	P3	TAA TCA GGA TTA AAT GGC CTT GGT ATG TTT GGC TCG GCT T	65.3
	P4	CAT TGC TGG ACT AAT TGC CAT CGT CAT GGT TAC AAT CTT G	63.7
	P5	GCT AAA GCT TTT AAG TCA TGC AAC AAA GCA AGA TTG TAA CCA TGA CGA	64.1
	P6	TGG CAA TTA GTC CAG CAA TGA AGC CGA GCC AAA CAT ACC A	67.7
	P7	AGG CCA TTT AAT CCT GAT TAG TTG	61.7

	P8	TTG TGT TAC ATA GGT T TGA AGG CTT TGA AGT CTG CCT GTA ATT AAC CTG TCA ATT T	63.2
B7.1	**5'**	GTC GAC GCT GAC TTC TCT ACC CCC AA	72.2
	3'	AAG CTT AAG GAA GAC GGT CTG TTC AGC	71.9
CEA w/o leader	**5'**	CTA GTG GTG AAG GGA GGA CAA CCT GGG AGA GGG TGG GAG GAG GGA GCT GGG GTC TCC TGG GTAGGA CAG GGC TGG	75.3
	3'	TCG ACC AGC CCT GTC CTA CCC AGG AGA CCC CAG CTC CCT CCT CCC ACC CTC TCC CAG GTT GTC CTC CCT TCA CCT	76.5
SARS m1	**5'**	TGG CTC GGC TTC ATT ATT GGA	56.6
	3'	TCC AAT AAT GAA GCC GAG CCA	56.6
SARS m2	**5'**	TGG CTC GGC ACC ATT ATT GGA	59.2
	3'	TCC AAT AAT GGT GCC GAG CCA	59.2
SARS m3	**5'**	TGG CCT TGG TAT GTT CCA CTC	57.1
	3'	GAG TGG AAC ATA CCA AGG CCA	57.1
IVH3H	**5'**	AGC TTT GCC CAA AGT ACG TGA AGC AAA ACA CAC TTA AAC TGG CTA CCG GAA TGA GAA ACG TGC CAG AAA AGC AAA CAT AAC	70.9
	3'	TCG AGT TAT GTT TGC TTT TCT GGC ACG TTT CTC ATT CCG GTA GCC AGT TTA AGT GTG TTT TGC TTC ACG TAC TTT GGG CAA	71.3
β-globin intron	**5'**	ACA GCT CCT GGG CAA CG	58.3
hGH poly(A)	**3'**	AAG GCT GGT GGG CAC TGG	61.0

2.1.2 Cell lines

CT26 (mouse colon cell, ATCC: CRL-2638)

PT67 (mouse retrovirus-packaging cell line, ATCC: CRL-12284)

Balb/3T3 (mouse embryo fibroblast, ATCC: CCL-163.2)

P338D1 (mouse lymphoblast, ATCC: TIB-39)

2.1.3 Bacterial strains

Escherichia coli Top10 strain: for general cloning (Invitrogen)

Escherichia coli DH5α™-T1: for site-directed mutagenesis (Invitrogen)

Salmonella typhimurium SL3261 strain: for plasmid uptake (kindly provided by Dr. Wu, Chang-Jer)

2.1.4 Plasmids

plasmid	Description	Source
pAAV-MCS	pCMV promoter for MCS	Stratagene
pAAV-CEA	EcoRI-XbaI fragment, containing the CEA fragment	This study
pAAV-CEA-SARS	EcoRI-HindIII fragment, containing the CEA-SARS fragment	This study
pAAV-CEA-m1	Modified pAAV-CEA-SARS, with one mutation site	This study
pAAV-CEA-m2	Modified pAAV-CEA-SARS, with two mutation sites	This study
pAAV-CEA-m3	Modified pAAV-CEA-SARS, with three mutation sites	This study
pAAV-B7.1	Modified pAAV, with a B7.1 fragment	Liao's lab
pAAV-CEA-B7.1-IVH3H	EcoRI-XhoI fragment, containing the CEA-B7.1-IVH3H fragment	This study
pMSCVneo	neomycin resistant and containing 5' LTR and a viral packaging signal	BD
pMSCVneo-CEA	EcoRI-XhoI fragment, containing the CEA fragment	This study

2.1.5 Chemicals, Enzymes, and reagents

Chemical	Source	Catalog number	Application
100bp DNA ladder	Protech	M1-100T	DNA electrophoresis
1kb DNA ladder	Protech	M1-1KB	DNA electrophoresis
3,3'-dioctadecyloxacarbocyanine （DIOC18）	SIGMA	D-4292	Cell staining
Agarose	MDBio	929049	DNA electrophoresis
Ampicillin	AMRESCO	0339	Bacterial culture
DMSO	MP	196055	Buffer
EDTA	Tedia	ER-0531	Cell passage
EtBr	AMRESCO	3434B14	DNA staining
Ethanol	SIGMA	E7023	DNA extraction
Fetal Bovine Serum	Biological industries	04-001-1A	Cell culture
HCl	Scharlau	AC0741	Buffer
Incomplete Freund's adjuvant	SIGMA	F-5506	Immunization
Isopropanol	C-Echo	PH-3101	DNA extraction
Kanamycin	MDBio	226039	Antibiotics
LB agar	AMRESCO	J637	Bacterial culture
LB broth	Scharlau	02-385	Bacterial culture
Lipofectamine 2000	Invitrogen	11668-019	Transfection

MEM	GIBCO	41500-034	Cell culture medium
NaCl	AMRESCO	0241	Buffer
NaHCO3	MP	194847	Additional ingredient to cell culture medium
NaOH	Showa	1943-0150	Buffer
Pfu polymerase	MDBio	826049	PCR
Propium iodide (PI)	SIGMA	P4170	Cell live/dead staining
PSA	Biological Industries	03-033-1B	Cell culture
RPMI	GIBCO	31800-022	Cell culture medium
Taq polymerase	BioKit	Bio Taq	PCR
Taq DNA polymerase XL	Protech	P6a	PCR
Trypan blue stain	GIBCO	0759	Cell staining
Trypsin	GIBCO	27250-018	Cell passage
Tween 20	MP	194724	ELISA

2.1.6 Antibodies

Antibodies	Source	Catalog number
Goat anti-mouse B7.1	R&D	AF740
Goat anti-mouse HRP	SIGMA	A0412
Goat anti-mouse FITC	Jackson ImmunoResearch	115-095-003

Mouse anti-mouse CEA	This study	N/A

2.1.7 Kits

Kit	Source	Catalog number	Application
Gene-Spin™ 1-4-3 DNA extraction kit	Protech	PT-DNA143XL-V2	DNA extraction, clean-up
Gene-Spin™ Miniprep Purification Kit	Protech	PT-MP530XLO-V2	DNA extraction
Gene Tailor™ Site-Direted Mutagenesis System	Invitrogen	12397-014	DNA site-directed mutagenesis
NucleoBond PC100	Macherey-Nagel	740573	DNA extraction
Mouse IL-2	DuoSet	DY402	ELISA
Mouse IL-4	DuoSet	DY404	ELISA
Mouse IL-10	DuoSet	DY417	ELISA
Mouse IL-12 p70	DuoSet	DY419	ELISA
Mouse IFN-γ	DuoSet	DY485	ELISA
Mouse TNF-α	DuoSet	DY410	ELISA
Mouse Th1/Th2 Cytokine	BD	551287	Cytometric bead array
SuperSignal West Pico Chemiluminescent Substrate	PIERCE	34080	Dot blot

2.1.8 Buffers

- 1X ACK lysis buffer

  0.15 M NH_4Cl, 1.0 mM $KHCO_3$, 0.1 mM Na_2EDTA in dd H_2O

- 1X PBS (pH7.4)

  137 mM NaCl, 10 mM Na_2HPO_4, 2.7 mM KCl, 1.8 mM KH_2PO_4

- 2% Blocking buffer

  1g nonfat powdered milk dissolved in 50ml 1X PBS buffer

- 50X TAE buffer

  48.4 g Tris base, 0.5 M EDTA (pH8.0) 20 ml, 11.42 ml acetic acid. dd H_2O was added to 200 ml.

- EDTA-trysin

  2.5g trypsin, 0.1 M EDTA (pH8.0) in 1L 1X PBS, pH7.4, 0.2 μm filtered

- Reagent Diluent

  1% BSA in PBS, pH7.2-7.4, 0.2 μm filtered

- Staining buffer

  1% BSA, 0.05%NaN_3 in 1x PBS

- Stop solution (for ELISA)

  1 N HCl

- PBST

  0.5% Tween 20 in 1x PBS

- Versene

  0.2g EDTA in 1L 1X PBS

- Wash buffer

  0.05% Tween 20 in PBS, pH7.2-7.4

2.1.9 Media

- LB (Luria-Bertani) broth

1% tryptone, 0.5% yeast extract, 1% NaCl

- LB (Luria-Bertani)/Ampicillin broth

  1% tryptone, 0.5% yeast extract, 1% NaCl, 50μg/ml ampicillin

- LB (Luria-Bertani)/Ampicillin agar

  1% tryptone, 0.5% yeast extract, 1% NaCl, 1.5% agar, 50μg/ml ampicillin

- DMEM medium

  10% FBS, 1% PSA in Dulbecco's Modified Eagle's Medium

- RPMI 1640 medium

  10% FBS, 1% PSA, 2g $NaHCO_3$ in 1L RPMI Medium 1640

- Opti-MEM I Medium

  Medium without serum

2.1.10 Equipment

-20℃ low temperature refrigerator (Frigidaire)

4℃ refrigerator (MINI KINGCON)

-80℃ low temperature refrigerator (NUAIRE)

Bench top orbital shaker 060 (LMS)

Biophotometer DPU-414 (eppendorf)

Bright-line chamber (Marienfeld)

Centrifuge 5415D (eppendorf)

Centrifuge 5804 R (eppendorf)

DNA electrophoresis unit Gel Mate 2000 (Toyobo)

Dot-blot machine (Bio-East)

Flow cytometer, FACSarray (BD)

Flow cytometer, FACScan (BD)

Heating block (FIRSTEK)

Inverted research microscope, IX71 (Olympus)

Laminar flow hood, Forma Class II, A 1284 (NSF)

Microplate reader, Sunrise (Tecan)

Microscope, CX31 (Olympus)

Orbital Shaking incubator OS1500R (TKS)

pH meter SP701 (Suntex)

Thermal cycler (eppendorf)

Uni-photo gel image system (EZ lab)

Water bath (FIRSTECK)

2.2 Methods

2.2.1 Computer prediction

The most probable epitope within the CEA-SARS sequence was calculated by the Internet software, SYFPEITHI (http://www.syfpeithi.de/). The sequence was pasted in the required box, and H2-Kd was selected as the MHC type. The number of amino acids of the epitope was set to nonamers (9 aa).

According to the description on the website, a reliability of at least 80% in retrieving the most apt epitope can be expected. Thus the naturally presented epitope should be among the top-scoring 2 % of all peptides predicted in 80 % of all predictions (http://www.syfpeithi.de/). The sequence that had the highest score was chosen to be our mutation template. To yield a mutated sequence, each amino acid of the template was replaced until another sequence had a highest score among all the changed sequences. It was named m1. Next, the amino acids within this new sequence was changed until another sequence with the highest score was

attained, which was named m2. The same approach was applied so that three cumulative mutations were generated in the m3 sequence.

2.2.2 PCR reaction

2.2.2.1 CEA synthesis

Because the antisense sequences overlap with the sense sequences for 20 bp in the designed primers, they will base pair with each other. Therefore, the whole CEA sequence can be constructed in the polymerase chain reaction (PCR) by DNA polymerase.

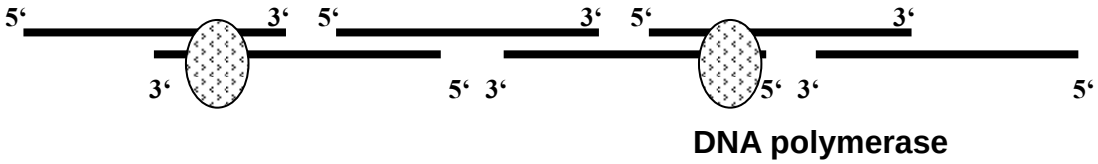


Figure 1. The scheme of CEA and SARS synthesis.

The CEA P1~P6 primers and the CEA P1~P5, P7 were added into a 0.2 ml tube, respectively, each taken 1 μl. The concentration of the first and the last primers in the sequence was 10 μM (P1, P6, P7) while the rest was 1 μM. 5 μl Taq buffer, 4 μl 10 nmol/ml dNTP, 0.5 μl Pro Taq, and 34.5 μl dd H_2O were added. The reaction cycle is: 94℃ for 30 sec, 50℃ for 30 sec, and 72℃ for 20 sec for 2 cycles, 94℃ for 30 sec, 60℃ for 30 sec, and 72℃ for 20 sec for 34 cycles, 72℃ for 5 min to complete the reaction.

2.2.2.2 SARS synthesis

The SARS P1~P8 primers were added into a 0.2 ml tube, each taken 1 μl. The concentration of the first and the last primers in the sequence was 10 μM (P1, P8) while the rest was 1 μM. 5 μl Taq buffer, 4 μl 10 nmol/ml dNTP, 0.5 μl Pro Taq, and 34.5 μl dd H_2O were added. The reaction cycle is: 94℃ for 30 sec, 55℃ for 30 sec, and 72℃ for 20 sec for 2 cycles, 94℃ for 30 sec, 65℃ for 30 sec, and 72℃ for 20 sec for 33 cycles, 72℃ for 5 min

to complete the reaction.

2.2.2.3 SARS mutation

The m1 plasmid was mutated by Gene Tailor^{TM} Site-Direted Mutagenesis System (Invitrogen, USA). The site-directed mutation has been designed on the primer beforehand. Original plasmid DNA is methylated before PCR so that it can be distinguished from the PCR product. After transformation into DH5α^{TM}-T1 competent cells, *Mcr*BC endonuclease in the host cell digests the methylated template DNA, leaving only unmethylated, mutated product.

100 ng plasmid DNA, 1.6 µl methylation buffer, 1.6 µl 10X SAM, 1.0 µl DNA methylase (4 U/µl) were added into a 0.2 ml tube, sterile, distilled water was added to make the total volume 16 µl. These reagents were incubated at 37℃ for 1 hr. Then, 2 µl methylated DNA, 5 µl 10X PCR buffer, 1.5 µl 10 mM dNTP, 1.5 µl SARS m1 5' and 3' primers (10 µM each), and 0.5 µl TaqXL (Protech, Taipei, Taiwan) were added into a 0.2 ml tube. The PCR condition is as the following: 94℃ for 2 min, 94℃ for 30 sec, 53℃ for 30 sec, 68℃ for 10 min (the last three steps were run for 24 cycles), and 68℃ for 10 min to finish the unfinished reaction.

The m2 and m3 plasmids were constructed by the following methods. The original CEA-SARS sequence was run twice to create two segments in which mutation sequence was designed in the primers. Then, these two segments served as the template for a second run of PCR, in which the very beginning and the end of the original CEA-SARS primers were added. After the construct of m2, m3 was created by the m2 template by the same process.

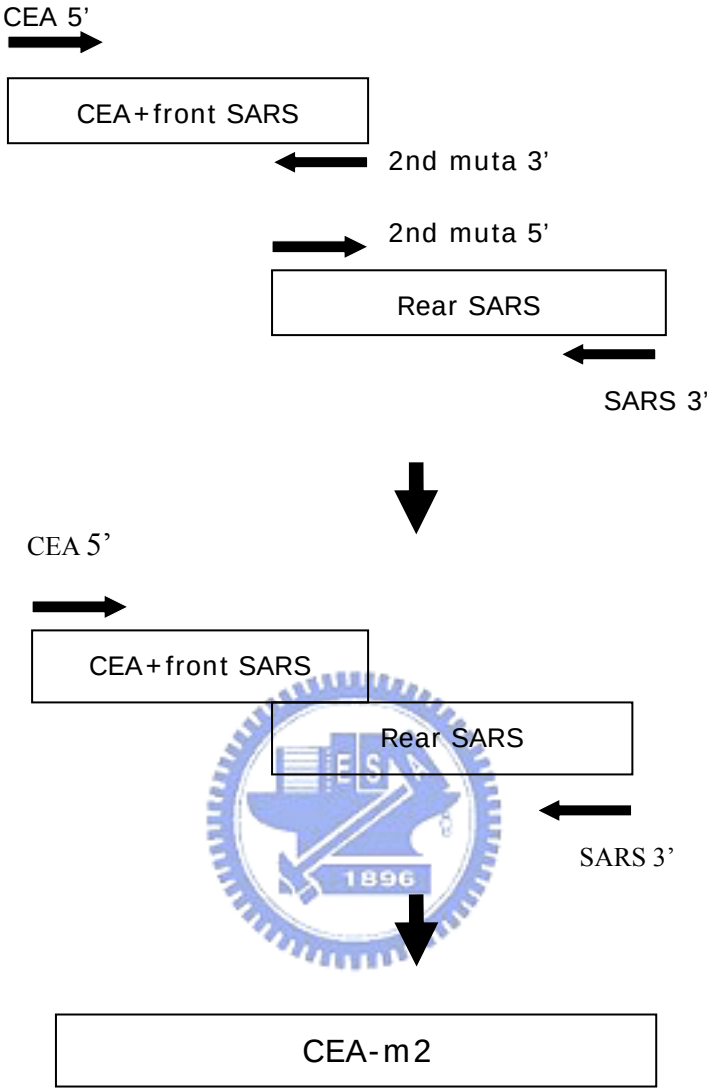


Figure 2. The scheme of m2, m3 plasmid construction.

The m1 plasmid serves as the template for m2 PCR. 5 μl template (~4.0 μg/ml), 5 μl 10X buffer, 4 μl 10 mM dNTP, 0.5 μl Pfu were added into a 0.2 ml tube. 1 μl CEA P1 and SARS m2 3' primers were added for the CEA+front SARS PCR while 1 μl SARS P8 and SARS m2 5' primers were added for the rear SARS PCR. To make the total volume be 50 μl, an appropriate amount of dd H_2O was added. The PCR condition for the CEA+front SARS sequence is as the following: 94℃ for 2 min, 94℃ for 30 sec, 55℃ for 30 sec (the last three steps were run for 34 cycles), 72℃ for 1 min, and 72℃ for 5 min. The PCR condition for the rear SARS sequence is: 94℃ for 2 min, 94℃ for 30 sec, 53℃ for 30 sec (the last three steps were run for 34 cycles), 72℃ for 1 min, and 72℃ for 5 min.

To combine the CEA+front SARS and the rear SARS sequences, 1 μl each PCR product was added, along with 5 μl 10X buffer, 4 μl dNTP, 0.5 μl Pfu, and 1 μl CEA P1 and SARS P8. The PCR condition is: 94℃ for 30 sec, 57℃ for 30 sec, 72℃ for 1 min. This step was run for 3 cycles. The annealing temperature was set at 57℃ for the annealing of the PCR products. Then, the next round of PCR condition is: 94℃ for 30 sec, 50℃ for 30 sec, 72℃ for 1 min. This step was run for 34 cycles. The annealing temperature was set at 50℃ for the annealing of the primers onto the templates.

The m2 plasmid serves as the template for m3 PCR. 3 μl template (~4.5μg/ml), 5 μl 10X buffer, 4 μl 10 mM dNTP, 0.5 μl Pfu were added into a 0.2 ml tube. 1 μl CEA P1 and SARS m3 3' primers were added for the CEA+front SARS PCR while 1 μl SARS P8 and SARS m3 5' primers were added for the rear SARS PCR. To make the total volume be 50 μl, an appropriate amount of dd H_2O was added. The PCR condition for the CEA+front SARS sequence is as the following: 94℃ for 2 min, 94℃ for 30 sec, 52℃ for 30 sec (the last three steps were run for 34 cycles), 72℃ for 1 min, and 72℃ for 5 min. The PCR condition for the rear SARS sequence is: 94℃ for 2 min, 94℃ for 30 sec, 48℃ for 30 sec (the last three steps were run for 34 cycles), 72℃ for 1 min, and 72℃ for 5 min.

To combine the CEA+front SARS and the rear SARS sequences, 1 μl each PCR product

was added, along with 5 μl 10X buffer, 4 μl dNTP, 0.5 μl Pfu, and 1 μl CEA P1 and SARS P8. The PCR condition is: 94℃ for 30 sec, 58℃ for 30 sec, 72℃ for 1 min. This step was run for 3 cycles. The annealing temperature was set at 58℃ for the annealing of the PCR products. Then, the next round of PCR condition is: 94℃ for 30 sec, 50℃ for 30 sec, 72℃ for 1 min. This step was run for 34 cycles. The annealing temperature was set at 50℃ for the annealing of the primers onto the templates.

2.2.3 Plasmid construction

2.2.3.1 Restriction enzyme digestion

0.5 μg~1 μg DNA was dissolved in an appropriate volume of water and was digested with restriction enzymes (following the commercial protocol). Generally, 1 μg DNA was digested with 5 unit of restriction enzyme in a 10 μl reaction at 37℃ overnight.

2.2.3.2 DNA extraction

The DNA digested by restriction enzyme was cleaned up by Gene-Spin™ 1-4-3 DNA extraction kit (Protech), following the commercial protocol.

The DNA solution was spun at 13,000 rpm for 30 sec in the spin column. The filtrate in the collection tube was discarded. 700 μl Washing solution (Protech Co., Taipei, Taiwan) was added and the solution was spun for 1 min at 13,000 rpm. This step was repeated twice. Then, the filtrate was discarded by centrifugation at 13,000 rpm for 3 min to remove residual trace of ethanol. The column was additionally incubated at 65℃ for 5 min to evaporate ethanol. DNA was eluted by 30-50 μl dd H_2O in a new tube.

2.2.3.3 Ligation

Generally, the concentration ratio of the vector and the insert was 1:3. The concentration of the insert and the vector was measured by a photometer. 1 μl 10 mM ATP, 1 μl T4 ligase,

1 μl 10X ligation buffer, and an appropriate volume of dd H_2O were added into a 500 μl tube to 10 μl. The mixture was incubated at 16℃ overnight.

2.2.4 Transformation of *E. coli*

2.2.4.1 Preparation of competent cells

One pick of *E. coli* was inoculated in 3 ml of LB broth and grew for 12 hr at 37℃ with vigorous shaking (~225 rpm). One ml of the overnight culture was transferred into 100 ml LB broth and was then incubated at 37℃ with shaking (~225 rpm) until the OD_{600} was between 0.35~0.45. The culture was set on ice for 10 min. The cells were recovered by centrifugation at 4,100 rpm for 10 min and then resuspended in 30 ml ice-cold 0.1 M $CaCl_2$. The cells were pelleted by centrifugation at 4,100 rpm for 10 min at 4℃. The pellet was resuspended in 2 ml 0.1 M $CaCl_2$ (containing 10% glycerol). The cells were dispensed at 100μl per eppendorf tube and then were stored at -80℃.

2.2.4.2 Transformation

Stored competent cells were thawed on ice. 1 μl~2 μl of plasmid DNA was mixed with 100 μl competent cells and was then stored on ice for 30 min. The mixture was incubated in a preheated 42℃ heating block for 90 sec and quickly thawed on ice for 2 min. Then 250 μl of LB broth was added to the cells. The culture was incubated at 37℃ with shaking (~225 rpm) for 50 min. 100 μl of the culture was plated on the LB agar plate with 50 μg/ml ampicillin. The plate was inverted and then incubated at 37℃ for 12~18 hr.

2.2.5 Plasmid DNA extraction

2.2.5.1 Minipreparation

Plasmid DNA in *E. coli* was extracted with Gene-Spin™ Miniprep Purification Kit (Protech). The procedure is as the following:

A single colony of *E. coli* was inoculated in 3 ml of LB broth (with antibiotics) and grew overnight at 37℃ with vigorous shaking (~225 rpm). One to two ml of the cells were recovered by centrifugation at 13,000 rpm for 1 min and then resuspended in 200 µl Solution I buffer (Protech Co., Taipei, Taiwan) in a new tube. 200 µl Solution II buffer (Protech Co., Taipei, Taiwan) was added and mixed gently. 200 µl Solution III buffer (Protech Co., Taipei, Taiwan) was added to the mixture and mixed gently again. Cells were spun at 13,000 rpm for 5 min at 4℃. The lysate was transferred to the Mini spin column. The solution was centrifuged at 13,000 rpm for 30 sec. The filtrate in the collection tube was discarded. 700 µl of Washing Solution (Protech Co., Taipei, Taiwan) was added in. The solution was spun at 13,000 rpm for 1 min. This step was repeated once again. After the filtrate was discarded, the column was centrifuged at 13,000 rpm for 3 min and incubated at 65℃ for 5 min to remove residual trace of ethanol. DNA was eluted by 30-50 µl dd H_2O and centrifuged at 13,000 rpm for 1 min. Plasmid DNA was stored at -20℃.

2.2.5.2 Midipreparation

One ml of precultured *E. coli* was added into 100 ml LB broth (with antibiotics) and incubated at 37℃ with shaking (~225 rpm) for 12-16 hr. The broth was centrifuged at 8,000 rpm at 4℃ for 15 min. After supernatant was discarded, 4 ml Buffer S1 (Macherey-Nagel, Inc., Duren, Germany) was added and the solution was vortexed to dispense the pellet. Then 4 ml of Buffer S2 (Macherey-Nagel, Inc., Duren, Germany) was added. The lysate was mixed gently by inverting the tube 6~8 times and incubated at room temperature for 2~3 min. The solution was mixed with the 4 ml pre-cooled Buffer S3 (Macherey-Nagel, Inc., Duren, Germany) and inverted gently 6~8 times until a homogeneous suspension containing an off-white flocculate was formed. The suspension was incubated on ice for 5 min. A NucleoBond AX 100 column was equilibrated with 2.5 ml Buffer N2 (Macherey-Nagel, Inc., Duren, Germany). The flow-through was emptied by gravity flow and discarded. The

bacterial lysate was cleared by centrifugation at 12,000 rpm at 4℃. The lysate was then loaded onto the NuceloBond column, which was emptied by gravity flow. Ten ml of Buffer S3 (Macherey-Nagel, Inc., Duren, Germany) was added to wash the column. This step was repeated once again. Plasmid DNA was eluted with 5 ml of Buffer N5 (Macherey-Nagel, Inc., Duren, Germany). Then 3.5 ml isopropanol was added to precipitate the eluted plasmid DNA. The mixture was incubated on ice for 10 min and centrifuged at 13,000 rpm for 30 min at 4℃. The supernatant was discarded. One ml 70% ethanol was added to the pellet and stored at -20℃ or the solution was centrifuged at 13,000 rpm for 5 min for further application. Last, the pellet was redissolved in 20 μl dd H_2O.

2.2.6 Cell culture

2.2.6.1 Balb/3T3

Balb/3T3 was cultured in DMEM (Sigma-Aldrich, St. Louis, USA) supplemented with 10% FBS and 1% PSA. Cells were incubated in tissue culture incubator with 5% CO_2 at 37℃.

2.2.6.2 PT67

PT67 was cultured in DMEM (Sigma-Aldrich, St. Louis, USA) supplemented with 10% FBS and 1% PSA. Cells were incubated in tissue culture incubator with 5% CO_2 at 37℃.

2.2.6.3 CT26

CT26 was cultured in RPMI 1640 (Invitrogen Co., USA) supplemented with 10% FBS, 0.2% $NaHCO_3$ and 1% PSA. Cells were incubated in tissue culture incubator with 5% CO_2 at 37℃.

2.2.6.4 P338D1

P338D1 was cultured in RPMI 1640 (Invitrogen Co., USA) supplemented with 10% FBS,

0.2% $NaHCO_3$ and 1% PSA. Cells were incubated in tissue culture incubator with 5% CO_2 at 37℃.

2.2.7 Transfection of mammalian cells

2.2.7.1 Seeding cells

The medium in the 75T flask (Corning, NY, USA) was discarded. Three ml of EDTA-trypsin was added and the flask was incubated at room temperature for 5 min or until cells were detached. 5 ml of medium was added to dilute EDTA-trypsin. The solution was centrifuged at 1,500 rpm for 5 min at 4℃. The supernatant was discarded. Cells were resuspended in 2 ml medium. Certain amount of cells was stained by Trypan blue and calculated by a bright-line chamber (Marienfeld, Germany). 2.5×10^5 cells were seeded in each well of a 6-well plate (Corning, NY, USA). Three ml of medium was added and the cells were maintained in the incubator with 5% CO_2 at 37℃ for 24 hr for further transfection.

2.2.7.2 Lipofectamine™ 2000 transfection

Cells were transfected with different plasmid DNA by Lipofectamine™ 2000 (Invitrogen, USA). The transfection procedure was as following.

DNA was diluted in 250 μl Opti-MEM I Medium (GIBCO, USA) and mixed gently. 10 μl of Lipofectamine™ 2000 was gently mixed with 250 μl Opti-MEM I medium and incubated for 5 min at room temperature. The diluted DNA was combined with the diluted Lipofectamine™ 2000 for 20 min at room temperature. The medium in the cells were discarded and cells were gently washed with Opti-MEM I medium twice. The 500 μl DNA-Lipofectamine™ 2000 mixture was added to 80-90% confluent cells. 500 μl of Opti-MEM I medium was added into each well gently and the cells were incubated at 37℃ in a CO_2 incubator for 12 hr. Two ml of growth medium (DMEM or RPMI) was added into each well and cells were incubated at 37℃ in a CO_2 incubator for 24-48 hr prior to the following assay.

2.2.8 Infection of mammalian cells

CT26 was plated 12-18 hr before infection in a 6-well plate at the cell density of 1 x 10^5 per well. DMEM from packaging cells, PT67, was collected. Equal amount of RPMI 1640 medium was added and the mixture was filtered through a 0.45-μm cellulose acetate or polysulfonic (low protein binding) filter. Three ml of the mixed medium was added into CT26, which was then incubated in a CO_2 incubator for 24 hr. New DMEM was added into PT67 and collected 24 hr later by the same procedure. Then it was added into CT26 to replace the mixed medium. Infection was carried out four times within 4 days.

2.2.9 Dot-blotting

2.2.9.1 Preparation of the CEA antibody

The pAAV-CEA-B7.1-IVH3H plasmid was constructed beforehand. It was transfected into Balb/3T3 by the transfection protocol described above. 48 hr later, cells were harvested and centrifuged at 1,500 rpm for 5 min. The supernatant was discarded and 600 μl 1x PBS was added to resuspend the cells. 600 μl of incomplete Freund's adjuvant (SIGMA, USA) was added to emulsify the mixture. Each mouse received 200 μl of the emulsified mixture. Blood was collected one week later and sera was collected by centrifugation at 4,000 rpm at 4℃ for 30 min. The inoculation of plasmid was carried out once a week for a month. Every transfectant had been examined by cytometer for the existence of the construct.

2.2.9.2 Confirmation of the CEA antibody

To test whether the sera contained the CEA antibody, Balb/3T3 cells were first transfected with pAAV-CEA-B7.1-IVH3H by the transfection protocol described above. 48 hr later, medium was discarded. 1 ml of versene was added and cells were incubated at 37℃ for 5 min. 1 ml DMEM was added to harvest the transfectants. The cells was recovered by

centrifugation at 1,500 rpm at 4℃ for 5 min. The supernatant was discarded. 5 ml of staining buffer was added to gently suspend the cells. Sera were added as the 1st antibody and the mixture was incubated on ice for 1 hr. The pellet was collected by centrifugation at 1,500 rpm at 4℃ for 5 min. The supernatant was discarded and washed with 1 ml staining buffer twice. Then, 2nd antibody was added into the cell solution and the cells were incubated in dark on ice for 30 min. The mixture was washed with 1 ml staining buffer and centrifuged at 1,500 rpm at 4℃ for 5 min. This step was repeated once. Finally, the cells were resuspended in 1 ml staining buffer and filtered by a mesh before further analysis by cytometer.

2.2.9.3 Dot-blotting

The supernatant of the infected CT26 was collected in a 1 ml tube. Cells were recovered by centrifugation at 1,500 rpm at 4℃ for 5 min and resuspended in 100 ml PBS. Cells were then lysed by repeated freeze-thaw cycles. Both kinds of samples (the supernatant and the lysed cells) were applied onto the nitrocellulose (NC) paper, which was prewetted with 1x PBS buffer on a dot-blot machine (Bio-East, Taiwan). Samples were vacuumed gently for 30 min. The NC paper was blocked by 2% blocking buffer for 30 min and washed with PBST three times (5 min, 10 min, and 10 min) at room temperature. The sera containing the CEA antibody were diluted 1000X in staining buffer and applied onto the NC paper gently at room temperature for 1 hr with shaking. The mixture was then discarded. The NC paper was washed with PBST three times (5 min, 10 min, and 10 min) at room temperature. The 2nd antibody conjugated with HRP was diluted 1000X in staining buffer and applied on the NC paper for 30 min in dark with shaking. Then, the mixture was discarded. The NC paper was washed with PBST three times (5 min, 10 min, and 10 min) at room temperature. The substrate was applied onto the NC paper for 5 min in dark. The NC paper was covered in the lead blocker (Okamoto, Japan) with the film for 25 min. Then, the film was developed in the developer for 1 min. The film was washed in water before it was stained in the fixer for 1

min.

2.2.10 Transformation of *Salmonella typhimurium*

2.2.10.1 Preparation of competent cells

A colony of *Salmonella typhimurium* was inoculated in 25 ml of LB and grew for 20 hr at 37℃ with vigorous shaking (~225 rpm). The overnight culture was transferred into 500 ml SOB containing 2 M $MgCl_2$ and was then incubated at 37℃ with shaking (~225 rpm) until the OD_{600} was between 0.35~0.4. The cells were recovered by centrifugation at 2,500 rpm for 15 min at 4℃ and the supernatant was discarded. An appropriate amount of dd H_2O was added to resuspend cells. This step was repeated twice. The pellet was resuspended in 1 ml dd H_2O with 10% glycerol. The cells were dispensed at 20 µl per eppendorf tube and then were stored at -80℃.

2.2.10.2 Transformation

Stored competent cells were thawed on ice. 1 µl of plasmid DNA was mixed with 20 µl competent cells and transferred into a pre-cooled cuvette. Cells was then electroporated at 2.5 mF, 2.5 kV, and 200Ω for 4~5 msec. The mixture was immediately recovered in 1 ml LB, transferred to a test tube, and incubated at 37℃ with agitation (~225 rpm) for 1 hr. 100µl of the culture was plated on the LB agar plate with 50 µg/ml ampicillin. The plate was inverted and then incubated at 37℃ for 12~18 hr.

2.2.11 P338D1 incubation with transformed *Salmonella typhimurium*

2.5 x 10^8 *Salmonella typhimurium* transformed with pAAV-hrGFP were incubated with 5 x 10^6 P338D1 seeded in the 10-cm dish for 2 hr in the 5 ml DMEM medium without PSA. Then, 5 ml of DMEM medium supplemented with 30 µg/ml kanamycin was added to kill the remaining bacteria in the medium. The mixture was incubated at 37℃ in a CO_2 incubator for

48 hr. It was then analyzed by cytometry.

2.2.12 Killing assay

2.2.12.1 Animal immunization

Six- to eight-week old female Balb/c mice were purchased from the National Laboratory Center and housed in a temperature- and light-controlled room (12L:12D) at the Animal Maintenance Facility of National Chiao Tung University. Mice were immunized with 1 x 10^8 attenuated *Salmonella typhimurium* (in 20μl PBS), which had been transformed into different vectors (pAAV-CEA, pAAV-CEA-SARS, pAAV-m1, pAAV-m2, pAAV-m3) three times in two weeks. The negative group was fed with 20 μl PBS alone.

2.2.12.2 Splenocyte isolation

Mice were sacrificed by dislocation and their spleens were quickly harvested in a laminar flow hood. Spleens were placed in a 280μm-pored mesh and chopped by scissors. 10 ml of DMEM was slowly added onto the mesh while spleens were being ground until the spleen tissue became white. Single cell suspension was collected in a Petri dish and recovered by centrifugation at 1,500 rpm at 4℃ for 5 min. Supernatant was discarded and 5 ml 1X ACK lysis buffer was added for 5 min at room temperature. 1X ACK buffer can lyse the red blood cells while leaving the rest of the lymphocytes and leucocytes. The mixture was then diluted by 10 ml of DMEM and cells were recovered by centrifugation at 1,500 rpm at 4℃ for 10 min. After the supernatant was discarded, the cells were rinsed by 10 ml DMEM once more. Finally, cells were resuspended in DMEM and underwent cell calculation by trypan blue exclusion. For the 100:1 killing ratio, cells were plated in a 24-well plate at 5 x 10^6 per well. For the 50:1 killing ratio, cells were plated in a 24-well plate at 2.5 x 10^6 per well. For the 25:1 killing ratio, cells were plated in a 24-well plate at 1.25 x 10^6 per well.

2.2.12.3 Target cell staining

Target cells, including CT26/CEA, CT26, and YAC-1, were passaged by EDTA-trypsin and the cell number was determined by trypan blue exclusion. 3 x 10^6 cells were suspended in RPMI 1640 and 30 μl 3,3'-dioctadecyloxacarbocyanine（DIOC18）（Sigma, MO,USA）was added for YAC-1, and 15 μl DIOC18 for CT26/CEA and CT26, respectively. Cells were then incubated at 37℃ in a CO_2 incubator for 16 hr. Before the killing assay, target cells were rinsed twice at 1,500 rpm at 4℃ for 5 min. Then, the cell number was determined by trypan blue exclusion and seeded at 5 x 10^5 per well in a 24-well plate.

2.2.12.4 The killing assay

Splenocytes and target cells were mixed at different ratios (Effector/Target ratio, E/T ratio=100/1, 50/1, 25/1, 12.5/1) in a 24-well plate. Then, cells were centrifuged at 1,500 rpm at 4℃ for 5 min and were incubated at 37℃ in a CO_2 incubator for 4 hr. 500 μg/ml propium iodide (Sigma, MO, USA) was added at 1/10 of the total volume. The cells were processed by FACScan flow cytometer (BD, NJ, USA) and analyzed by CellQuest softeware.

2.2.13 The cytokine profile assay

2.2.13.1 The in vitro cytokine assay

The sera were analysed by ELISA kits, detecting IL-2, IL-12, IL-4, IL-10, IFN-γ, and TNF-α, by the commercial protocol as provided below:

The Capture Antibody (R&D, MN, USA) was diluted to the working concentration in PBS without carrier protein and coated immediately on a 96-well microplate (Disposable non-sterile assay plate. Corning, NY, USA) at 100 μl per well. The plate was sealed and incubated overnight at room temperature. Each well was aspirated and washed with 400 μl Wash Buffer (R&D, MN, USA) by a squirt bottle, manifold dispenser or autowasher. This step was repeated twice for a total of three washes. Plates were blocked by adding 300 μl of

Reagent Diluent (R&D, MN, USA) to each well and they were incubated at room temperature for 1 hr. The aspiration/wash step was repeated before sample addition. 100 µl samples or standard in Reagent Diluent were added. The plate was incubated at room temperature for 2 hr. The aspiration/wash step was repeated. 100 µl of the Detection Antibody (R&D, MN, USA) diluted in Reagent Diluent was added to each well. The plate was incubated at room temperature for 2 hr. The aspiration/wash step was repeated. 100 µl of the working dilution of Streptavidin-HRP was added to each well. The plate was incubated for 20 min at room temperature. The aspiration/wash step was repeated. 100 µl of Substrate Solution (R&D, MN, USA) was added to each well, and the plate was incubated for 20 min at room temperature. 50 µl of stop solution was added to each well and the plate was gently tapped to ensure thorough mixing. The optical density was determined by a microplate reader (Tecan) set to 450 nm and the data were analysed by Magellan5 software.

2.2.13.2 The in vivo cytokine assay

Mouse Th1/Th2 cytokine standards were reconstituted in Assay Diluent (BD, NJ, USA) for 15 min. Standards were serially diluted using the Assay Diluent. 10 µl of each mouse cytokine Capture Bead Suspension (BD, NJ, USA), including IL-2, IL-4, IL-5, IFN-γ, TNF-α, was mixed for one test. 50 µl of mixed beads was transferred to each assay tube. 50 µl/tube Standard Dilutions (BD, NJ, USA) and 25 µl/tube samples were added into the appropriate tubes. 50 µl/test PE Detection Reagent was added and the tubes were incubated in dark for 2 hr at room temperature. Samples were then washed with 1 ml Wash Buffer (BD, NJ, USA) and recovered by centrifugation at 200g for 5 min. Tubes were carefully aspirated and the supernatant was discarded. 300 µl of Wash Buffer was added to each assay tube to resuspend the bead pellet. Standards and samples were transferred to a 96-well plate and analyzed by FACSarray cytometer (BD, NJ, USA) and BD^{TM} CBA Software.

2.2.14 Tumor inoculation

2.2.14.1 The protection assay

Six- to eight-week old Balb/c mice (N=5) were immunized by the protocol in 2.2.11.1. One week after the last boost, mice were inoculated with 5 x 10^5 CT26/CEA, which was passaged by versene.

2.2.14.2 Therapy assay

Balb/c mice (N=3) were inoculated with 1 x 10^5 CT26/CEA (in 200 μl PBS/each mouse) and were orally immunized with 1 x 10^8 *Salmonella typhimurium* (in 20 μl PBS/each mouse) four days later. They were re-immunized once a week after the first immunization.

2.2.15 Data analysis

Results are expressed as mean ± SE. Student's t test was applied to compare treatment effects in different groups. A value of $p < 0.05$ was considered significant.

Chapter 3 Results

3.1 Epitope prediction (CEA-SARS) by Internet software

3.1.1 The epitope score of the CEA-SARS sequence

The CEA-SARS sequence was calculated by the Internet software, SYFPEITHI. The score of the whole sequence is listed in **Table 1**. The results showed that the epitopes in the CEA sequence got low scores. The highest score within the CEA sequence is 16. The sequence with the highest score (scored 21) was chosen to be our epitope target, which was within the SARS sequence and would be mutated later (WYVWLGFIA). The SARS sequence, therefore, would be a better immunogen than the CEA sequence. In addition, the fusion of CEA and SARS sequences did not create a new epitope with a higher score.

3.1.2 Affinity of H2-Kd compared with other epitopes

To determine the immunogenicity of the CEA sequence, the binding affinity between the H2-Kd and the CEA sequence was calculated first by the Internet software as a rough estimation. According to the description in the software, the maximal scores vary between different MHC alleles (Hans-Georg Rammensee 1999). It is clear that the CEA sequence is less immunogenous by the affinity calculation because its top 5 highest epitopes score far lower than the well-known epitopes (LLO_{91-99}, $p60_{217-225}$, $EGFP_{200-208}$) (**Table 2**). Secondly, the epitope within the CEA-SARS sequence was also compared with these known epitopes in Listeria monocytogenes (LLO_{91-99}, $p60_{217-225}$) and $EGFP_{200-208}$ (**Table 2**). The score of the CEA-SARS epitope is not as high as those of the well-known epitopes. As a result, we mutated this epitope in hope that mutation would enhance the immunogenicity of our DNA vaccine.

3.2 Point mutation prediction (CEA-SARS) by Internet software

The epitope within the SARS sequence, WYVWLGFIA, served as a template and was mutated by the Internet software, SYFPEITHI. Every amino acid within this epitope was substituted so as to yield an epitope with the highest score. The flow chart of the mutation process is presented in Table 3. The original SARS epitope, WYVWLGFIA, scores 21 whereas after amino acid substitution, the highest score is 25 (**Table 4**). There are two such epitopes, WYVWLGFIL and WYVWLGFII (Table 4), the last of which was selected to be the first mutated sequence (named m1).

This m1 epitope further served as the template for the next round of mutation. Every amino acid within this epitope was also substituted so as to yield an epitope with the highest score among all the substitutions, WYVWLGTII, which scores 29 (**Table 5**). It was named m2, which has accumulated two mutations within the original SARS eptiope. The same procedure was applied to yield m3, WYVPLGTII, which scores 32 and contains three cumulative mutations within (**Table 5**).

3.3 Construction of pAAV-CEA, pAAV-CEA-SARS, pAAV-CEA-m1, pAAV-CEA-m2, pAAV-CEA-m3 expression plasmids

The CEA sequence was obtained by polymerase chain reaction (PCR) amplification with primers CEA P1~P6 (with the stop codon) or with primers CEA P1~P5, P7 (without the stop codon). The PCR mixture and the PCR condition are described in 2.2.2.1. Both PCR products were digested with restriction enzyme EcoRI at the 5' end and XbaI at the 3' end. The restricted fragment was then ligated with the large EcoRI-XbaI fragment of pAAV-MCS. Both constructs were transformed into *E. coli* and cells were cultured for 12-16 hr. The PCR

with β–globin intron and hGH poly(A) primers proves that the CEA sequence with a stop codon was constructed into the plasmids (**Figure 3**). The CEA sequence without a stop codon was confirmed by the restriction enzyme EcoRI at the 5' end and XbaI at the 3' end into a 135 bp fragment after mini-prep DNA extraction (**Figure 4**).

To construct pAAV-CEA-SARS, the SARS sequence was first obtained by PCR amplification with primers SARS P1~P8. The PCR mixture and the PCR condition are described in 2.2.2.2. The SARS PCR product was digested with restriction enzyme XbaI at the 5' end and HindIII at the 3' end. The restricted fragment was then ligated with the large XbaI-HindIII fragment of pAAV-CEA (without the stop codon). The construct was transformed into E. coli and cells were cultured for 12-16 hr. Then, plasmid DNA was extracted by minipreparation (described in 2.2.5.1) and digested by XbaI and HindIII into 174 bp fragment to confirm the existence of the insert (**Figure 5**). The plasmid was named pAAV-CEA-SARS.

To construct pAAV-CEA-m1, the m1 sequence was obtained by PCR amplification with primers SARS m1 5' and SARS m1 3' by Gene Tailor^{TM} Site-Direted Mutagenesis System (Invitrogen, USA), described in 2.2.2.3. The m1 plasmid was transformed into DH5α™-T1 and cells were cultured for 12-16 hr. After plasmid DNA extraction by minipreparation (in 2.2.5.1), the plasmid was digested by XbaI and HindIII into a 174 bp fragment to confirm the existence of the insert (**Figure 6**). The plasmid was named pAAV-CEA-m1.

To construct pAAV-CEA-m2, the m2 sequence was obtained by PCR amplification with primers CEA P1, SARS m2 5', SARS m2 3', and SARS P8, which is described in 2.2.2.3. The PCR product was digested with restriction enzyme EcoRI at the 5' end and HindIII at the 3' end. The restricted fragment was then ligated with the large EcoRI-HindIII fragment of pAAV-MCS. The construct was transformed into *E. coli* and cells were cultured for 12-16 hr. After mini-preparation, the plasmid was digested by EcoRI and XhoIII into a 309 bp fragment to confirm the existence of the insert (**Figure 7**). The plasmid was named

pAAV-CEA-m2.

To construct pAAV-m3, the m3 sequence was obtained by PCR amplification with primers CEA P1, SARS m3 5', SARS m3 3', and SARS P8, which is described in 2.2.2.3. The PCR product was digested with restriction enzyme EcoRI at the 5' end and HindIII at the 3' end. The restricted fragment was then ligated with the large EcoRI-HindIII fragment of pAAV-MCS. The construct was transformed into *E. coli* and cells were cultured for 12-16 hr. After mini-preparation, the plasmid was digested by EcoRI and XhoIII into a 309 bp fragment to confirm the existence of the insert (**Figure 8**). The plasmid was named pAAV-CEA-m3.

The diagrams of these five constructs are shown in **Figure 9**.

3.4 Construction of pAAV-CEA-B7.1-IVH3H expression plasmid

The CEA sequence (without its leader sequence) was obtained by primer annealing at 95℃ (**Figure 10**). The annealing product was cooled down to room temperature before it was ligated with the large XbaI-SalI fragment of pAAV-B7.1 (kindly provided by Hsieh, Yuan-Ting). The construct was transformed into *E. coli* and cells were cultured for 12-16 hr. The PCR with β–globin intron and hGH poly(A) primers shows that a fragment of the predicted size (830 bp) was obtained (**Figure 11**). After seuquenced, the CEA sequence without the leader sequence was proved to have been constructed into the plasmids. It was named pAAV-CEA-B7.1. Then, the plasmid DNA was extracted by Midipreparation (described in 2.2.5.2) for the IVH3H insert.

The IVH3H (^{305}CPKYVKQNTLKIATGMRNVPEKQT^{328}) represents the carboxyl-terminal 24 residues of the influenza virus H3 subtype hemagglutinin (HA) heavy chain (HA_1) and has been proved to be able to enhance Th2 pathway (Fahrer, Geysen et al. 1995). Thus, IVH3H was constructed into the plasmid to induce antibody secretion. The construct was also obtained by primer annealing at 95℃. The 5' protruding end had been

designed as the digested HindIII site, and the 3' protruding end as the digested XhoI site. The insert was ligated with the large HindI-XhoI fragment of pAAV-CEA-B7.1. The construct was transformed into *E. coli* and cells were cultured for 12-16 hr. The PCR with primers β–globin intron and IVH3H 3' shows that a fragment of the predicted size (870 bp) was obtained (**Figure 12**). After sequenced, the CEA-B7.1-IVH3H sequence was proved to have been constructed into the plasmids.

The diagram of CEA-B7.1-IVH3H is shown in **Figure 13**.

3.5 Construction of pMSCVneo-CEA expression plasmid

The insert was obtained from pAAV-CEA (with a stop codon) by restriction enzyme digestion at EcoRI and XhoI sites at the 5' end and the 3' end, respectively. To exclude the ligation between the large EcoRI-XhoI fragment from pMSCVneo and the large EcoRI-XhoI fragment from pAAV-CEA, the latter was further digested by BglII at the 5' end and ClaI at the 3' end. After ligation (Vector/Insert=1/10), the plasmid was transformed into E. coli and bacteria were cultured for 12-16 hr. The restriction digestion by EcoRI and XhoI produced a 135 bp fragment in clone #1 and #2, which reveals the existence of the insert (**Figure 14**).

The diagram of pMSCVneo-CEA is shown in **Figure 15**.

3.6 Verification of plasmid expression by cytometry

To confirm that plasmids (pAAV-CEA, pAA-CEA-SARS, pAAV-CEA-m1, pAAV-CEA-m2, pAAV-CEA, m3) could be expressed by macrophages, pAAV-hrGFP (kindly provided by Chuang, Huai-Yao) was transformed into *Salmonella typhimurium* first by the protocol described in 2.2.10.2. hrGFP was inserted into the multiple cloning site (MCS), which was easy to be detected by fluorescence and substituted for other constructs. The overnight culture was refreshed with 7 ml LB (with ampicillin). 2.5 x 10^8 *Salmonella* were cultured with 5 x 10^6 P338D1, a macrophage-like cell line, as described in 2.2.11. The

expression of hrGFP was detected by cytometer (**Figure 16**). When compared with its negative counterpart, P338D1 uptaking pAAV-hrGFP expressed more fluorescence in FL1 (**Figure 16**).

3.7 CT26 infection by retrovirus: small fragment of CEA is secreted from CT26

3.7.1 Transfection of PT67 cell lines with pMSCVneo-CEA expression plasmid

For the CEA peptide to work as a less immunogenous TAA in the tumor model, it was constructed into CT26 by retroviral infection. The retroviral gene transfer technology is based on the coordinated design of packaging cell lines and retroviral expression vectors. pMSCVneo-CEA contains a packaging signal (Ψ) so that once in the packaging cell line (PT67), RNA from the vector is packaged into infectious, replication-incompetent retroviral particles.

Prior to infection, PT67 was transfected with pMSCVneo-CEA. Because the plasmid also contains neomycin resistant gene, G418, the analog of neomycin, was used for drug-resistance selection in the eukaryotic cells. According to the datasheet, PT67 underwent selection at 500μg/ml G418. On day 9, transfected PT67 was still confluent in the 75T flask (Corning, NY, USA) while the untransfected PT67 diminished quickly and only a little had remained (**Figure 17**).

3.7.2 Infection of CT26 with supernatant of PT67 transfection

24 hr after PT67 transfection, PT67 supernatant was collected. The infection procedure is described in 2.2.8. After four times of infection, CT26 was transferred into a 24-well plate for G418 selection. Prior to the selection, the drug-resistance test of CT26 had been carried out. 1.5 x 10^5 CT26 were seeded in a 6-well plate and were cultured for 24 hr. Different

concentrations of G418 were added into each well. On day 7, the cell condition in each well was distinguishable. CT26 all died at 300 μg/ml G418 and above (**Figure 18d, e, f**). Therefore, 400 μg/ml G418 was chosen as the CT26 selection concentration (**Figure 18**). The infected CT26 became gradually confluent in the well seven days after G418 selection (**Figure 19**).

3.7.3 The CEA expression by dot blot

To confirm that the CEA antigen is expressed by the eukaryotic cell line in our model, the supernatant of CT26 and lysed CT26 were both tested by dot blot. The procedure has been depicted in 2.2.9.3. The CEA antibody was first verified by transfecting Balb/3T3 with pAAV-CEA-B7.1-IVH3H and pAAV-B7.1. The pAAV-CEA-B7.1-IVH3H transfected Balb/3T3 had a higher expression of fluorescence than its negative control and pAAV-B7.1 transfected counterpart after surface marker staining by sera, indicating the sera contained the CEA antibody (**Figure 20**). Next, CT26 were transfected with pAAV-CEA-B7.1-IVH3H and were incubated in a CO_2 incubator at 37℃ for 48 hr. The supernatant and lysed CT26 cells were then collected. Dot blot reveals that CEA was expressed into the supernatant but not in the cell lysis (**Figure 21**).

3.8 In vitro killing assay of Balb/c splenocytes

3.8.1 The effect of DIOC18 staining on target cells

The effect of DIOC18 staining on target cells were tested before the killing assay because staining may affect the cell survival rate. The staining process was depicted in 2.2.11.3. The overnight staining on CT26/CEA, CT26, and YAC-1 revealed that DIOC18 not only stably expressed on the cell membrane after 24 hr but also contributed little to the cell mortality (**Figure 22**, **Figure 23**, **Figure 24**). When compared with their unstained counterparts, staining contributed 3.98% mortality to CT26/CEA, 1.17% to CT26, and 1.57%

to YAC-1 (**Figure 22**, **Figure 23**, **Figure 24**)

3.8.2 The killing assay

To analyze the killing efficiency of our constructs quantitatively, splenocytes in each immunized group (N=3) of mice were collected and were mixed with target cells, CT26/CEA, CT26, YAC-1, at different ratios (E/T ratio=100/1, 50/1, 25/1). The procedure was described in 2.2.11.4. Each group immunized with different plasmids (in *Salmonella typhimurium*) was later labeled as CEA (pAAV-CEA immunization), SARS (pAAV-CEA-SARS immunization), m1 (pAAV-CEA-m1 immunization), m2 (pAAV-CEA-m2 immunization), and m3 (pAAV-CEA-m3) groups.

At an E/T ratio of 100/1, CEA alone could not sufficiently enhance the immunity when compared with its negative control (Negative:15.00 ± 1.96%; CEA: 13.75 ± 1.45%). However, immunization both with CEA and SARS fragments, no matter mutated or not, could slightly enhance the killing efficiency (**Figure 25a**). Mutations in the m2 and m3 groups resulted in a significant difference when compared with the negative control and the CEA group (**Figure 25a**). The killing efficiency with mutated sequences were conspicuously improved at an E/T ratio of 50/1 (m1: 20.20 ± 3.55%; m2: 18.04 ± 2.60 %; m3: 22.08 ± 3.69%), as shown in **Figure 25b**. The m3 mutations resulted in a significant difference when compared with the CEA group. Such a phenomenon was also observed at the E/T ratio at 25/1, where the mutated groups, especially the m3 mutation, enhanced the killing of CT26/CEA (**Figure 25c**), indicating that mutations ensure the killing specificity to the CEA molecule. Beside the m3 mutation, there was also a significant difference between the m2 group and the negative control (12.81 ± 1.42% ; m2: 21.15 ± 3.77%).

The CT26 killing did not vary significantly among each group at an E/T ratio of 50/1, or 25/1 (**Figure 26b, c**). However, the m3 group had a significant difference when compared with the negative group at the E/T ratio of 100/1 (**Figure 26a**).

The killing percentage between CT26/CEA and CT26 was compared to analyze the specific killing to CT26/CEA. At an E/T ratio of 25/1, the specific killing to CT26/CEA in the CEA group did not enhance at all. On the contrary, it was aggravated when compared to the negative group (**Figure 27**). The situation was relieved in the SARS group and the specific killing to CT26/CEA was improved in the mutated groups. The specific killing is proportional to the number of mutations. In other words, the m3 group had the highest specific killing (8.93%), the m2 group ranked the second (3.74%), and the m1 last (0.82%), as shown in **Figure 27**.

To understand whether such an immunization could induce innate immunity, YAC-1, a MHC-less cell line, was targeted for NK cells. The CEA group could not induce innate immunity and showed no significance compared to the negative control at E/T ratios of 100/1, 50/1, and 25/1 (**Figure 28**). The SARS, m1, m2, and m3 groups, however, could slightly enhance the innate immunity at these ratios (**Figure 28**). There was a significant difference between the SARS group and the negative group, and between the m2 group and the negative group at an E/T ratio of 100/1 (**Figure 28a**). The m1 group had a significant difference when compared with the negative and the CEA group at the E/T ratio of 50/1 while the m3 group had a significant difference compared with the negative group (**Figure 28b**). At an E/T ratio of 25/1, the m2 group had a significant difference when compared with the negative and the CEA groups (**Figure 28c**).

3.9 Cytokine profile assay

3.9.1 The in vitro cytokine assay

In order to understand the cytokine profile among differently immunized groups quantitatively, 2 x 10^6 /well splenocytes in each group were stimulated by either CT26/CEA soup (specific stimulation) or CT26 (non-specific stimulation) soup for 24 hr. The cytokine

profile, including Th1 cytokines (TNF-α, IFN-γ, IL-2, IL-12) and Th2 cytokines (IL-4, IL-10), was detected by ELISA described in 2.2.13.

When compared with the negative group, the CEA group did not show any significant increase in TNF-α secretion when stimulated by CT26/CEA soup, but there was a significant decrease when stimulated by CT26 soup (**Figure 29**). The SARS, m1, m2, and m3 groups all showed a significant increase in TNF-α secretion under CT26/CEA soup stimulation (**Figure 29a**). The mutation groups, namely m1, m2, and m3, had a significant increase of TNF-α when compared with the CEA group (**Figure 29a**). Under CT26 soup stimulation, the SARS and m2 resulted in a significant increase of TNF-α compared with the negative group, but the m3 group resulted in a significant decrease of TNF-α (**Figure 29b**). When compared with the CEA group, the modification in the SARS, m1, and m2 increased significantly (**Figure 29b**).

IL-10 secretion was repressed significantly in the CEA group under CT26 soup stimulation (Neg: 22.12 ± 1.38 pg/ml; CEA: 13.21 ± 3.73 pg/ml). However, specific stimulation, the CT26/CEA soup, could slightly enhance its secretion in the CEA group (Neg: 14.44 ± 4.09 pg/ml; CEA: 23.20 ± 1.34 pg/ml), as shown in **Figure 30**. The SARS fragment, no matter mutated or not, could all strongly enhance IL-10 under both stimulations when compared with the negative and the CEA groups (**Figure 30**).

The CEA group secreted more IL-4 when compared with the negative control under both CT26/CEA soup (Neg: 3.33 ± 0.29%; CEA: 6.36 ± 3.40%) and CT26 soup stimulations (Neg: 2.60 ± 1.74%; CEA: 5.35 ± 2.56%). Yet IL-4 secretion was down-regulated in the SARS, m1, m2, and m3 groups (**Figure 31**). There was a significant decrease in the SARS and the m3 groups under the CT26/CEA soup stimulation (**Figure 31a**).

Compared with the negative control, the IL-12 secretion was slightly enhanced in the CEA group under CT26/CEA soup stimulation (Neg: 5.83 ± 0.64 pg/ml; CEA: 8.05 ± 0.60 pg/ml), but it was not significantly different under CT26 soup stimulation (**Figure 32**). The IL-12 secretion was suppressed in the SARS group under both conditions, but the

phenomenon was significant under CT26/CEA soup stimulation (**Figure 32a**). When compared with the CEA group, it also reveals a significant decrease (CEA: 8.05 ± 0.60%; SARS: 3.19 ± 0.60%). However, the IL-12 secretions were non-detectable under CT26/CEA soup stimulation in the m1 and m3 groups (**Figure 32a**). Though there was no significant decrease in IL-12 secretion in the SARS group when compared with the negative and the CEA groups under CT26 stimulation, the m1 and m2 groups were significantly down-regulated when compared with the negative group (**Figure 32b**). Moreover, groups with a mutation modification (m1, m2, and m3) resulted in a significant decrease in IL-12 when compared with the CEA group (**Figure 32b**).

The IFN-γ secretion was significantly elevated in the CEA group under CT26/CEA soup stimulation when compared with the negative group whereas there were no differences among other groups (**Figure 33a**). The CEA group was down-regulated under CT26 stimulations, while the m3 group had a significant elevation in IFN-γ secretion (**Figure 33b**) when compared with the negative and the CEA groups. IL-2 was all undetectable in our model (data not shown).

3.9.2 The in vivo cytokine assay

To understand the cytokine profile within the animal model, the in vivo cytokine assay was performed with Mouse Th1/Th2 Cytokine CBA (BD, NJ, USA). The procedure was described in 2.2.13.2. The result reveals that both Th1 cytokines, including IL-2, IFN-γ, TNF-α, and Th2 cytokines, including IL-4 and IL-5, were not enhanced in vivo in the CEA group when compared with the negative control. However, the SARS, m1, m2, and m3 groups could generally enhanced the Th1 and Th2 cytokine expression in vivo, suggesting that the SARS fragment, no matter mutated or not, could induce immune cells to secrete more cytokine against tumor (**Figure 34**).

3.10 Tumor growth

3.10.1 The protection assay

Mice were orally immunized by 1 x 10^8 *Salmonella typhimurium*. Then, 5 x 10^5 CT26/CEA cells per mouse were inoculated by the protocol in 2.2.14.1. The tumor volume was measured once every 2 or 3 days. Tumor volume was calculated with the formula: volume = length x width x height.

The tumor-free rate was highest in the m1 group (75%). Mice in the m3 group had a 60% tumor-free rate (**Figure 35**). Tumor volume was smaller in the SARS-modified groups and significantly decreased in the SARS and the m3 groups on day 26, indicating the SARS fragment could provide sufficient protection against the development of tumor, CT26/CEA, and the modifications on the SARS fragment ensures such a protective effect (**Figure 36**). After one month of tumor inoculation, m1 and m3 groups still had the highest survival rate (100%) while the negative group had 50% survival rate. The m2 group had the lowest survival rate (20%) on day 32 (**Figure 37**).

3.10.2 The therapy assay

Mice were inoculated with 1 x 10^5 CT26/CEA and four days later they were orally immunized with 1 x 10^8 *Salmonella typhimurium* transformed with pAAV-CEA and pAAV-SARS in 20 μl PBS, respectively. The negative group was fed with 20 μl PBS. Every week mice were re-immunized and tumor volume was measured once every two or three days. Mice in the SARS group had the smallest tumor volume, indicating the SARS fragment provided enough protection against CT26/CEA (**Figure 38**). The tumor volume was significantly different between the SARS group and the CEA group on day 44 (**Figure 38**).

H2-Kd nonamers		go to top
Position	1 2 3 4 5 6 7 8 9	**score**
78 (within SARS)	W **Y** V W L G F I **A**	21
59 (within SARS)	R **L** Q S L Q T Y **V**	19
89 (within SARS)	I **A** I V M V T I **L**	18
83 (within SARS)	G **F** I A G L I A **I**	17
5 (within CEA)	S **A** P P H R W C **I**	16
36 (within CEA)	W **G** L L G R T G **L**	16
42 (within SARS)	T **G** L K V E A E **V**	16
65 (within SARS)	T **Y** V T Q Q L I **R**	16

Table 1. **The epitope scores of the CEA-SARS sequence calculated by the Internet software, SYFPEITHI.** Epitopes that score 16 and above are listed. Epitope starting at the 78th amino acid in the CEA-SARS sequence has the highest score at 21, which serves as the template for affinity mutation.

H2-Kd nonamers		go to top	References	Affinity Strength
Position	1 2 3 4 5 6 7 8 9	**score**		
78	W **Y** V W L G F I **A**	21	CEA-SARS	Medium
5	S **A** P P H R W C **I**	16	CEA	Weak
36	W **G** L L G R T G **L**	16	CEA	Weak
11	W **C** I P W Q R L **L**	15	CEA	Weak
31	V **G** G G S W G L **L**	15	CEA	Weak
Listeria LLO_{91-99}	G **Y** K D G N E Y **I**	24	Nakamura, Y., et al. Infection and Immunity. 2003. 71(4), 1748–1754	Good
Listeria $p60_{217-225}$	K **Y** G V S V Q D **I**	27	H. G. Archie Bouwer et al. Infection and Immunity. 1996. 64(7), 2515-2522	Good
$EGFP_{200-208}$	H **Y** L S T Q S A **L**	27	Gambotto A, et al. Gene Therapy. 2001. 8(23). 1814-15.	Good

Table 2. **The CEA and CEA-SARS epitopes compared with other known epitopes that have been proved to elicit immunity in Balb/c mice.** The score of the CEA and CEA-SARS sequences are calculated by the Internet software, SYFPEITHI. The score of CEA is lower than the well-known epitopes of *Listeria* (LLO_{91-99} and $p60_{217-225}$) and $EGFP_{200-208}$. The epitope score of CEA-SARS is closer to these well-known epitopes and therefore, may be expected to elicit the immune response.

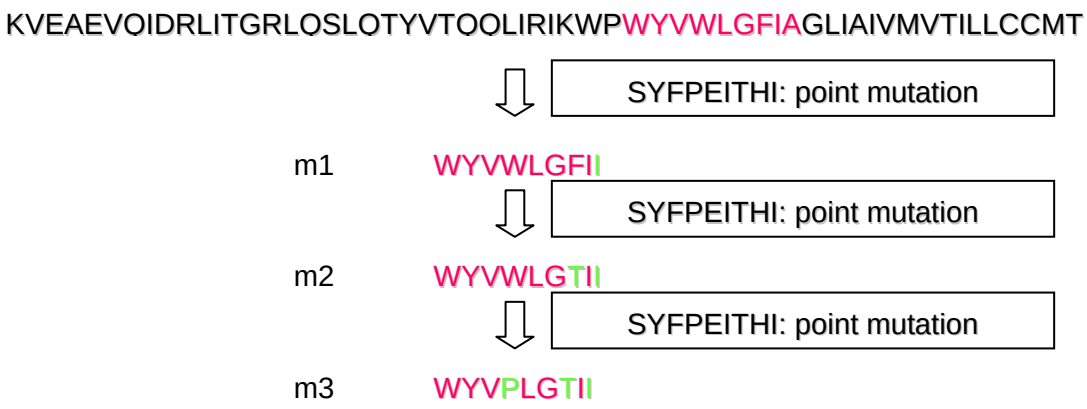


Table 3. The flow chart of mutagenesis. First, the epitope sequence is predicted. Then, every amino acid within the predicted epitope is substituted to gain higher affinity between the MHC molecule (H2-Kd) and the epitope. Among all, the epitope that has the highest score was chosen to be our first mutation construct and served as the template for the next round of mutation. It was named m1. The same approach was applied to obtain the two accumulative mutations in the m2 construct and three accumulative mutations in the m3 construct.

H2-Kd nonamers		go to top
Pos	1 2 3 4 5 6 7 8 9	**score**
(Template) 34	W **Y** V W L G F I **A**	21
1	W Y V W L G F E A	22
1	W Y V W L G F I V	23
1	W Y V W L G F I L	25
(m1) 1	W Y V W L G F I I	25

Table 4. The candidate sequence of the m1 construct. The epitope (WYVWLGFIA) in the CEA-SARS sequence was substituted by the approach described in the text. After substitution, two epitopes scored highest (at 25) were randomly chosen to be the first mutation construct (m1) and served as the template for the next round of mutation.

m2	1	W Y V W L G T I I	29
m3	1	W Y V P L G T I I	32

Table 5. The generation of m2 and m3. The construct of m2 (WYVWLGTII) was mutated from the m1 and the m3 (WYVPLGTII) construct was mutated from m2.

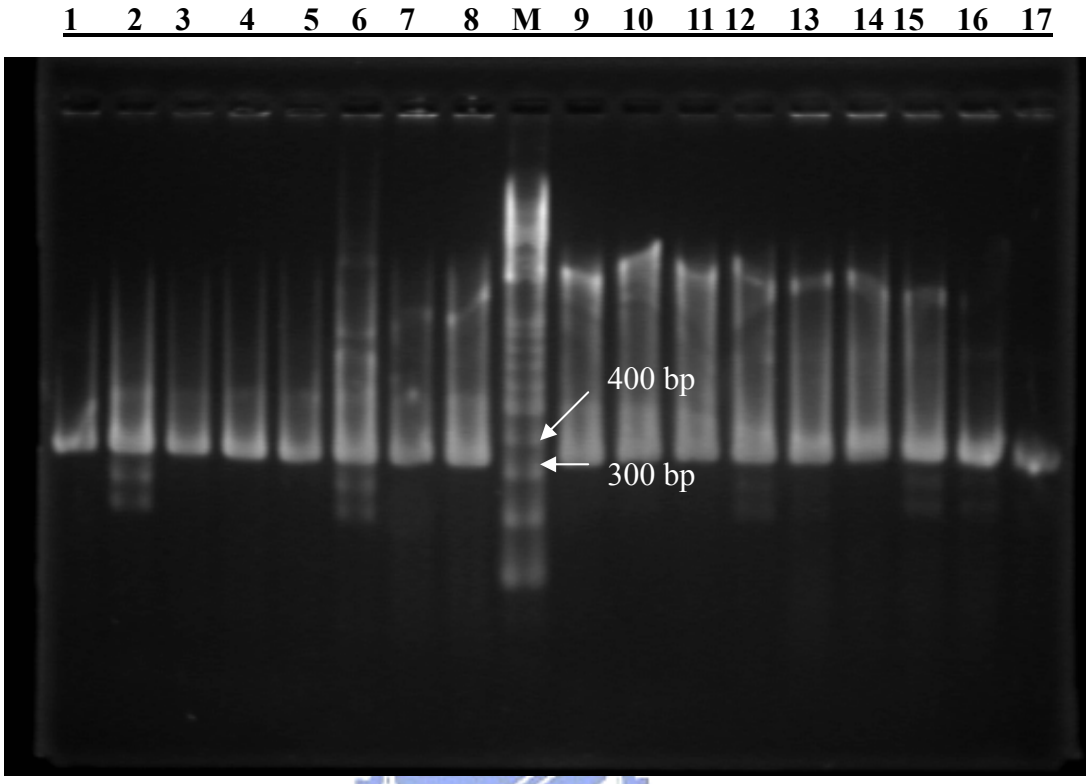


Figure 3. The construct of pAAV-CEA (with a stop codon). The CEA sequence with a stop codon was cloned into pAAV-MCS. The PCR with β–globin intron and hGH poly(A) primers proves that the CEA sequence with a stop codon was constructed into the plasmids (336 bp). Clone #1, #3, #4, and #5 were picked and sequenced. Clone #1 was chosen for plasmid DNA extraction. M: 100 bp ladder marker. #1~#17: clones.

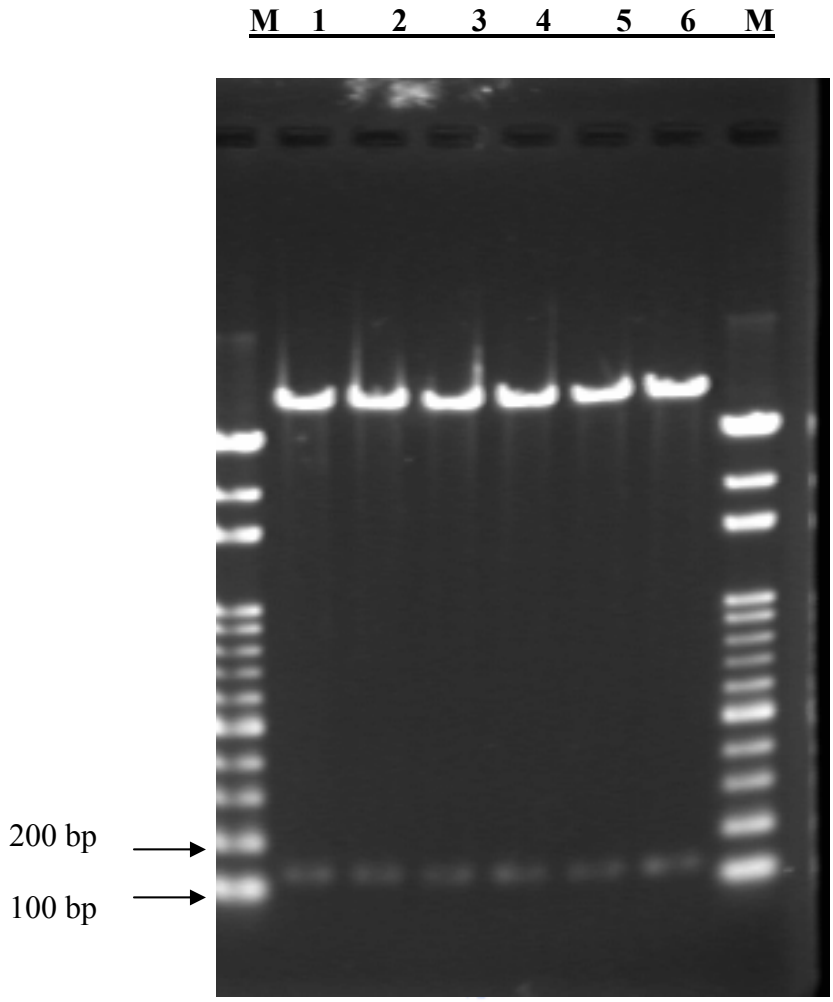


Figure 4. Restriction enzyme digestion of the pAAV-CEA construct without a stop codon. The plasmid was digested by EcoRI and XbaI into a 135 bp fragment. Clones #1, #2, and #3 were picked and sequenced. Clone #1 was chosen for plasmid DNA extraction. M: 100 bp ladder marker; #1~#6: clones.

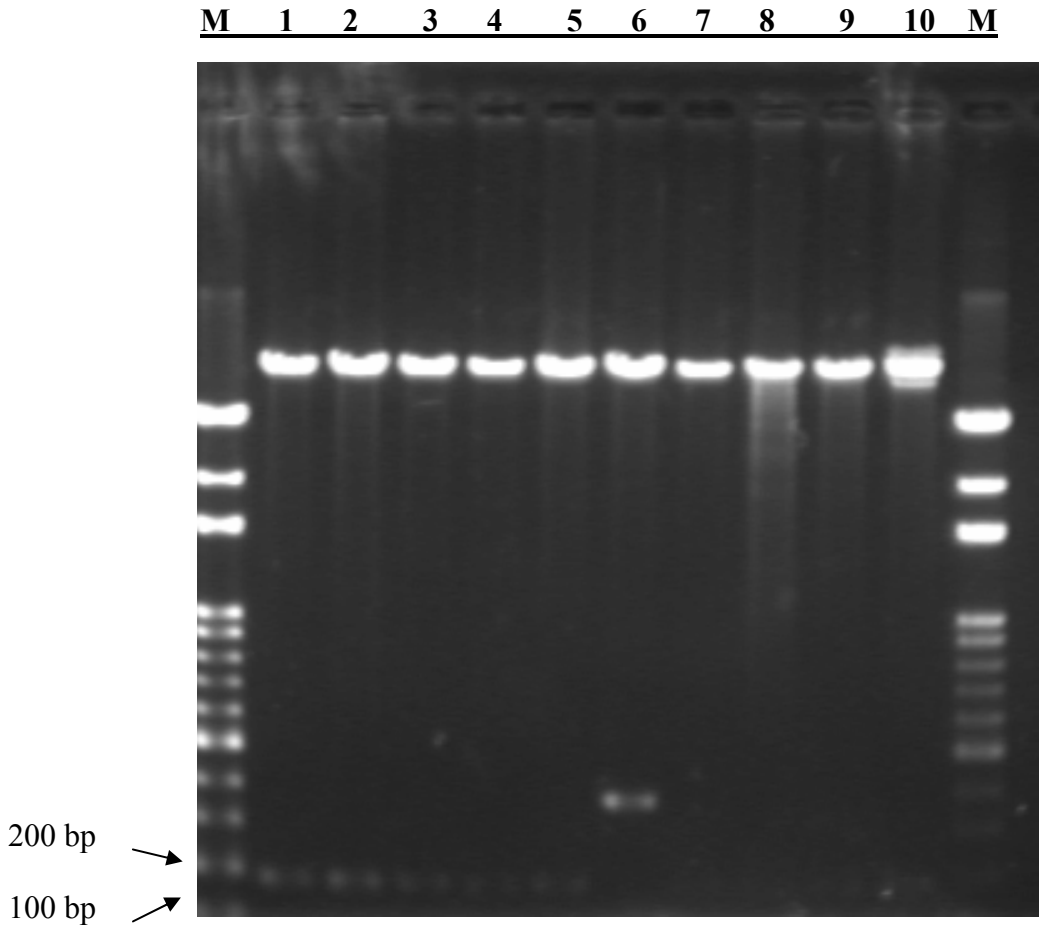


Figure 5. Restriction enzyme digestion of the pAAV-CEA-SARS construct. The plasmid DNA was digested by XbaI and HindIII into a 174 bp fragment. Clone #1, #2, and #3 were picked and sequenced. #2 was picked. M: 100 bp ladder marker; #1~#10: clones.

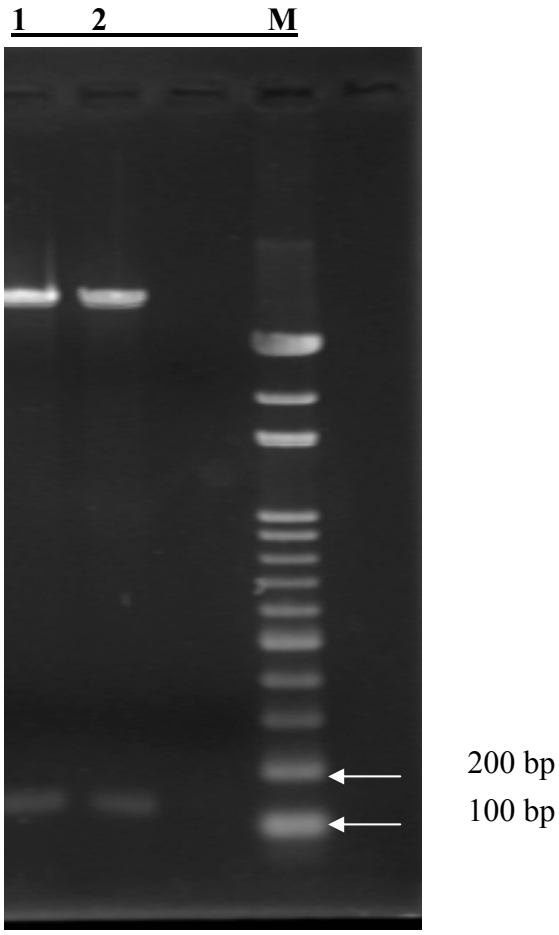


Figure 6. Restriction enzyme digestion of the pAAV-CEA-m1 construct. The plasmid was digested by XbaI and HindIII into a 174 bp fragment. Clone #1 and #2 were picked and sequenced. #2 was picked. M: 100 bp ladder marker; #1~#2: clones.

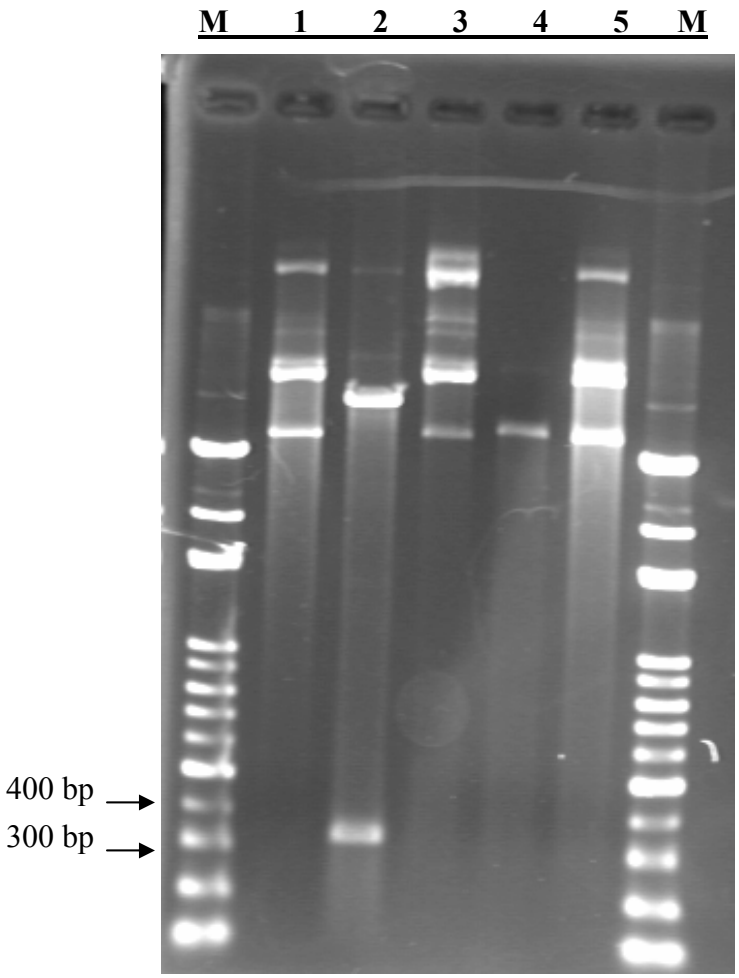


Figure 7. Restriction enzyme digestion of the pAAV-CEA-m2 construct. The plasmid was digested by EcoRI and XhoIII into a 309 bp fragment. Clone #2 picked and sequenced. M: 100 bp ladder marker; #1~#5: clones.

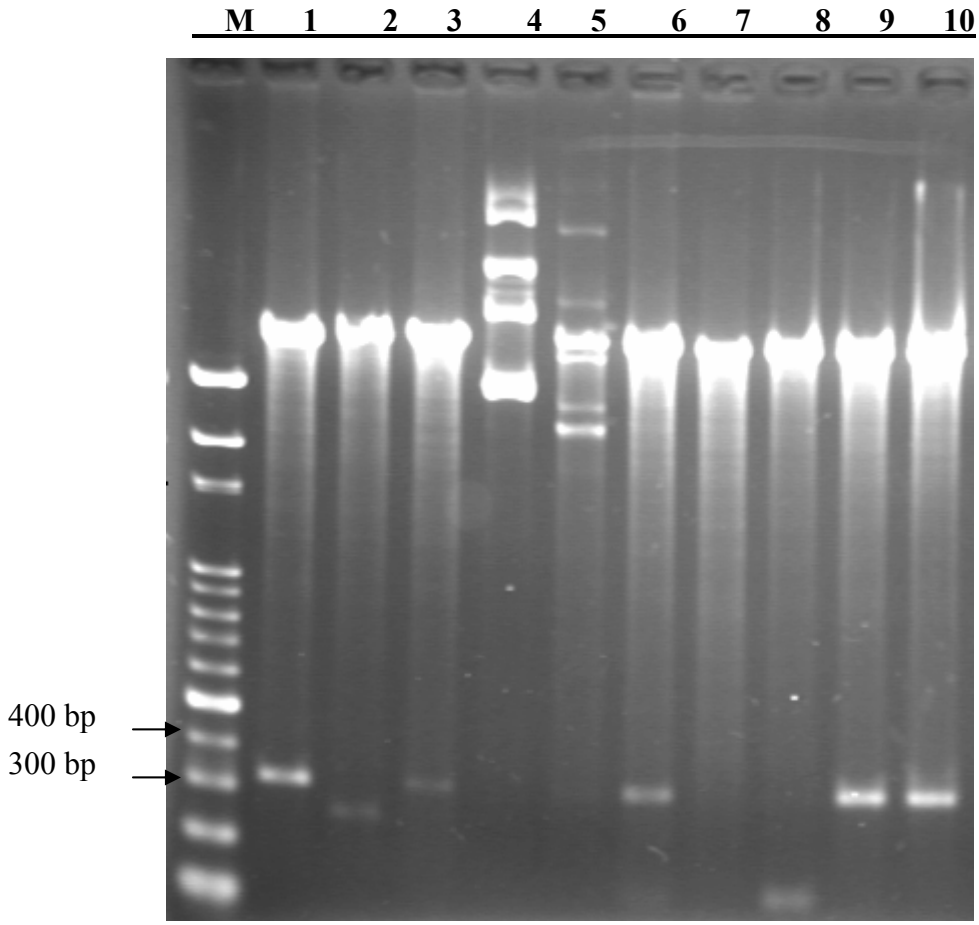


Figure 8. Restriction enzyme digestion of the pAAV-CEA-m3 construct. The plasmid was digested by EcoRI and XhoIII into a 309 bp fragment. Clone #9 and #10 were picked and sequenced. M: 100 bp ladder marker; #1~#10: clones.

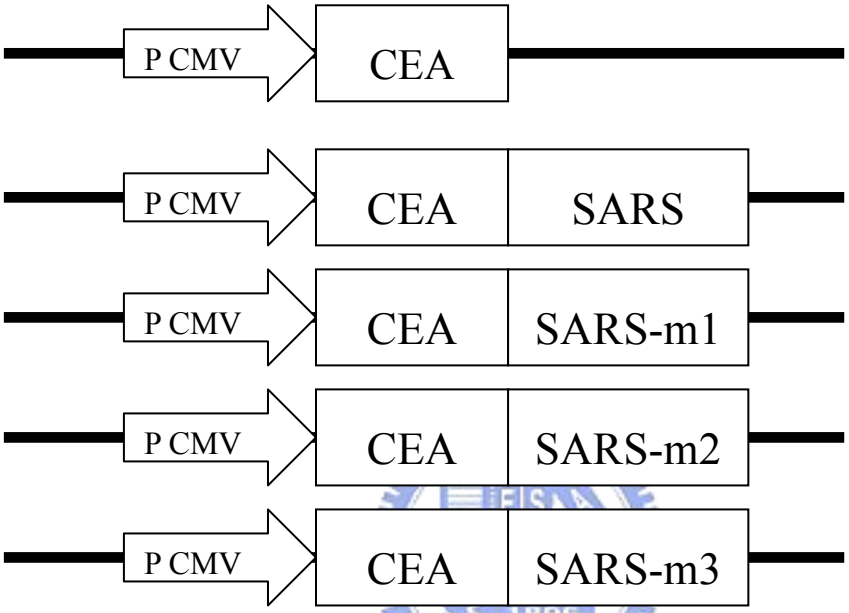


Figure 9. The diagram of pAAV-CEA, pAAV-CEA-SARS, pAAV-CEA-m1, pAAV-CEA-m2, and pAAV-CEA-m3 for immunization.

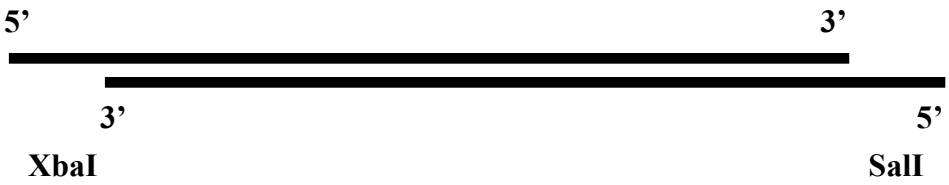


Figure 10. Primer annealing of the CEA sequence without its leader sequence. The primers were annealed at 95℃ and was cooled down to room temperature. The 5’ and 3’ end of primers have been designed as the stick end of *Xba*I (CTAGT) and *Sal*I (TCGAC) restriction sites. Therefore, the annealing product did not undergo PCR before ligation.

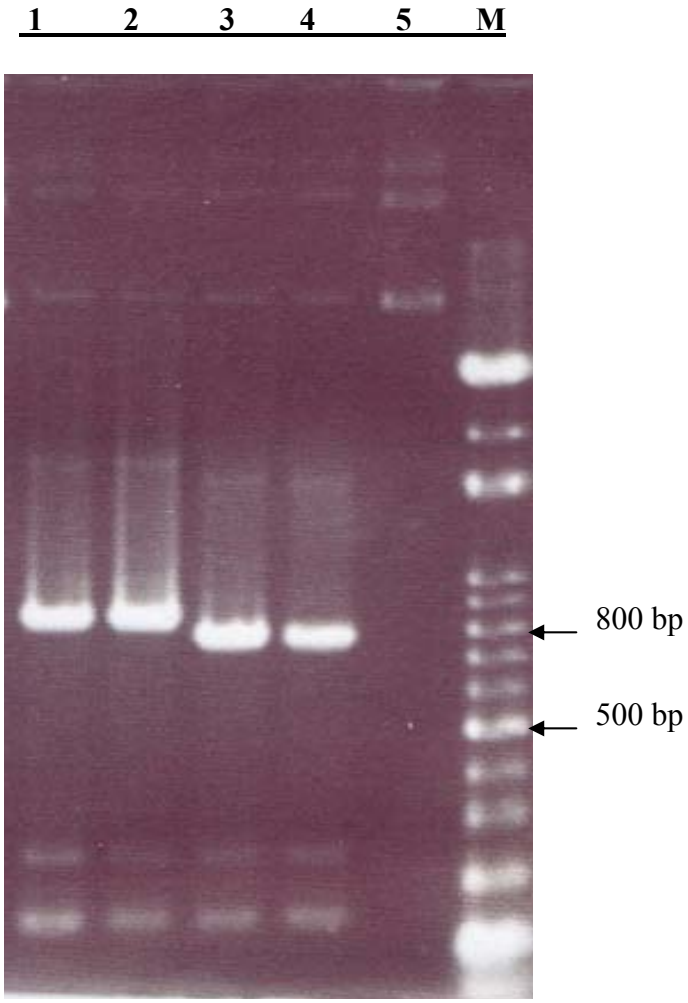


Figure 11. The construct of pAAV-CEA-B7.1. The CEA sequence without its leader sequence was cloned into pAAV-B7.1 The PCR with β–globin intron and hGH poly(A) primers shows that a fragment with the predicted size (830 bp) was obtained. #1 and #2 were picked and sequenced to prove the existence of the CEA sequence without its leader sequence. M: 100 bp ladder marker; #1~#5: clones.

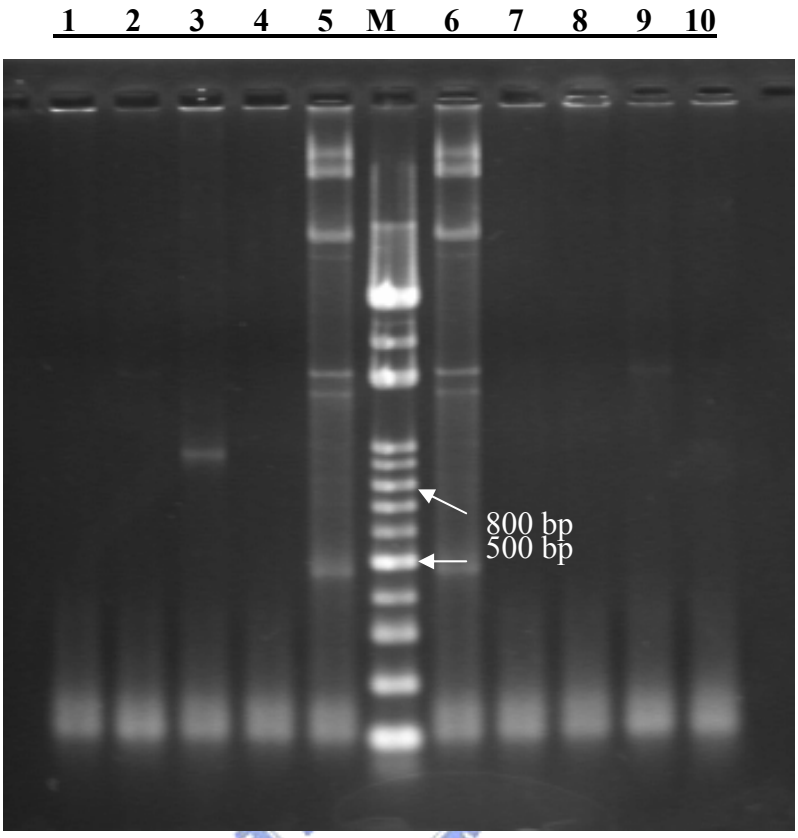


Figure 12. The construct of pAAV-CEA-B7.1-IVH3H. IVH3H was obtained from primer annealing and was cloned into pAAV-CEA-B7.1. The PCR with primers β–globin intron and IVH3H 3' shows that t that a fragment with the predicted size (870 bp) was obtained. Clone #3 was picked and sequenced to prove the existence of the CEA-B7.1-IVH3H sequence.

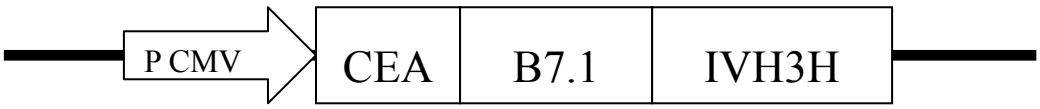


Figure 13. The diagram of pAAV-CEA-B7.1-IVH3H for CEA antibody production.

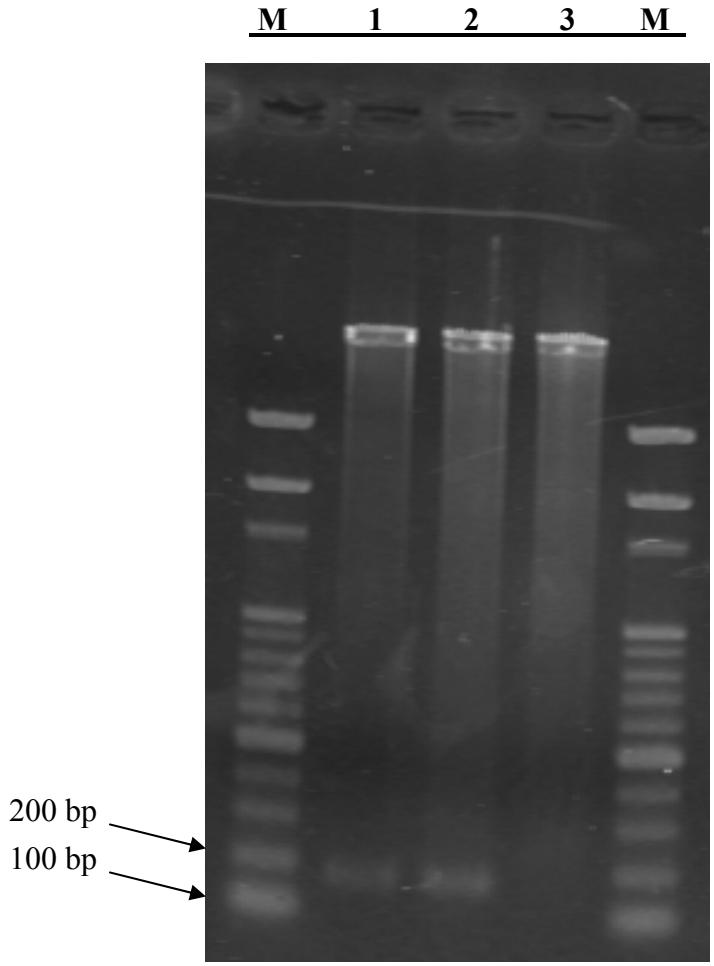


Figure 14. Restriction enzyme digestion of pMSCVneo-CEA. The restriction digestion by EcoRI and XhoI produced a 135 bp fragment in clone #1 and #2, the latter of which was picked and sequenced. M: 100 bp ladder marker; #1~#3: clones.

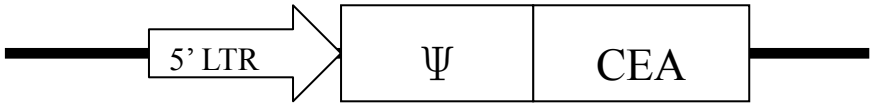


Figure 15. The diagram of pMSCVneo-CEA for CT26 infection.

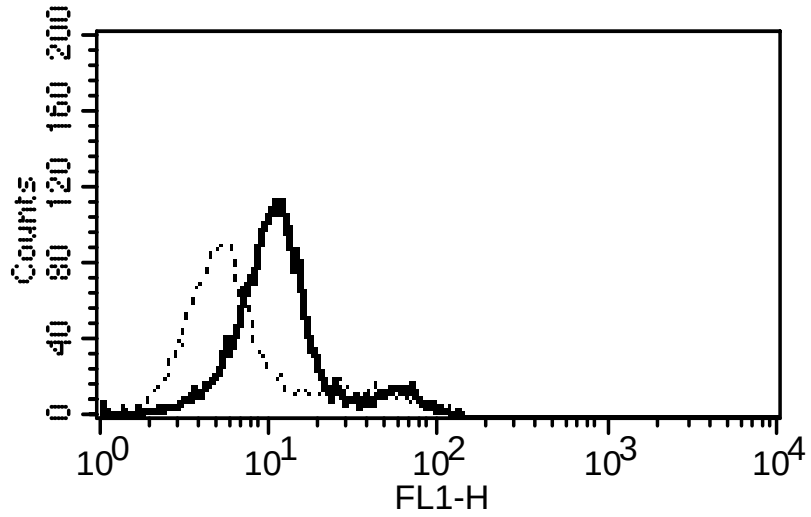
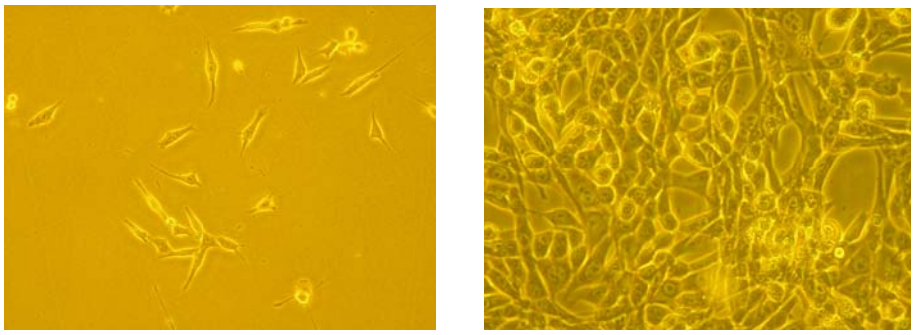


Figure 16. The fluorescence expression of hrGFP in P338D1. pAAV-hrGFP was transformed in *Salmonella typhimurium* and bacteria were incubated with P338D1 overnight. The expression of hrGFP was detected by cytometry (FL1). The pAAV-hrGFP transfected bacteria (dark line) had a higher fluorescence expression than its negative counterpart (dotted line). Dotted line: negative control (P338D1 + Salmonella w/o any plasmid); Dark line: P338D1 + Salmonella w/ plasmid.



a) **b)**

Figure 17. Selection of PT67 transfected with pMSCVneo-CEA. PT67 was selected at 500 μg/ml G418 on Day 9. a) Neg, PT67 without pMSCVneo-CEA transfection. Most of the PT67 cells had been killed. b) PT67 with pMSCVneo-CEA transfection. The transfected PT67 cells became confluent even under drug-resistance selection.

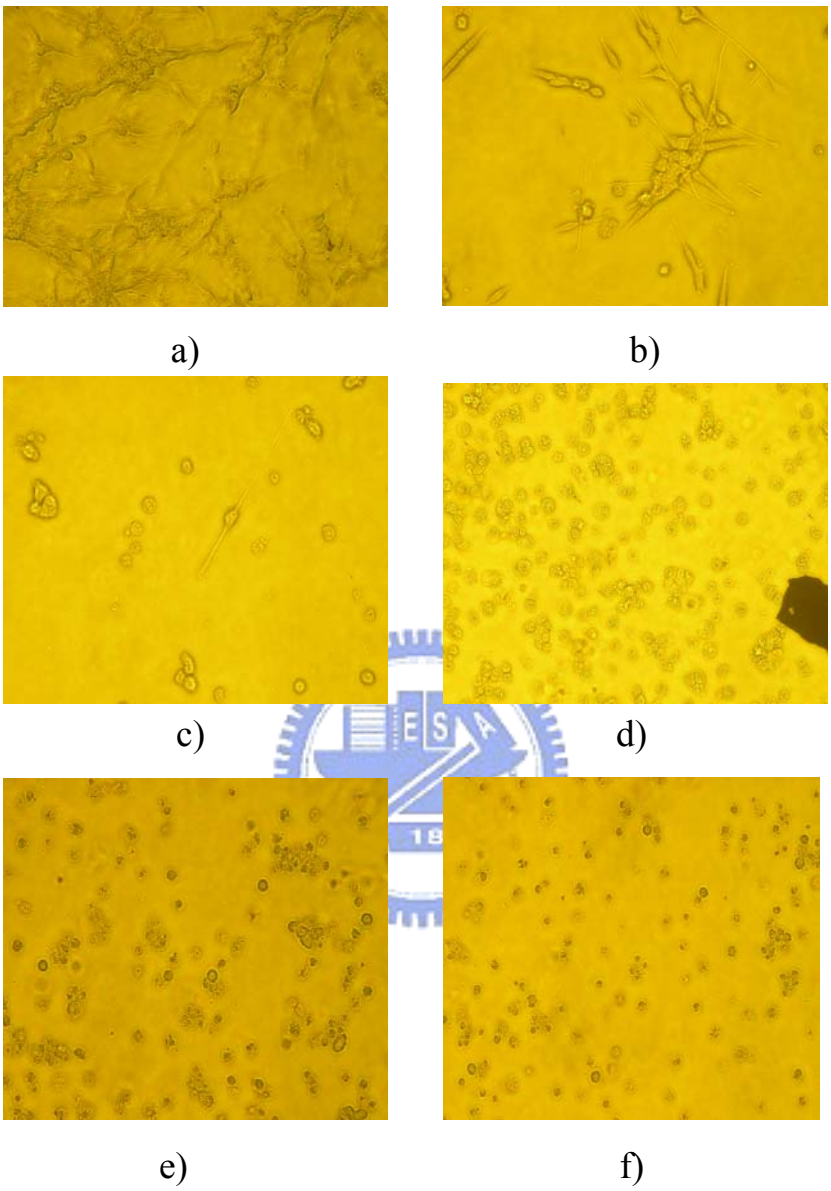
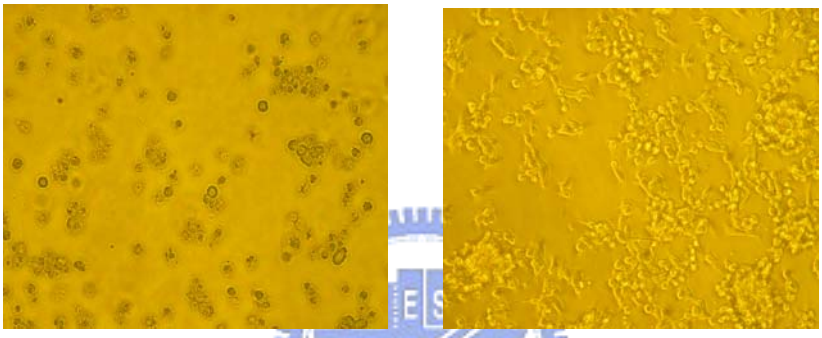
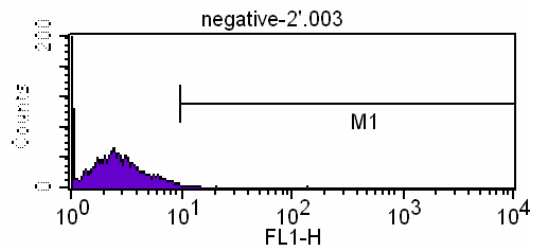


Figure 18. The G418 resistance test of CT26. Before infection, the G418 resistance test was performed at different concentrations. 400 μg/ml G418 was chosen for infection. On day 7, the cell condition at each concentration is shown. a) 0 μg/ml, confluent b) 100 μg/ml, partially dead c) 200 μg/ml, mostly dead d) 300 μg/ml, all dead e) 400 μg/ml, all dead f) 500 μg/ml, all dead.

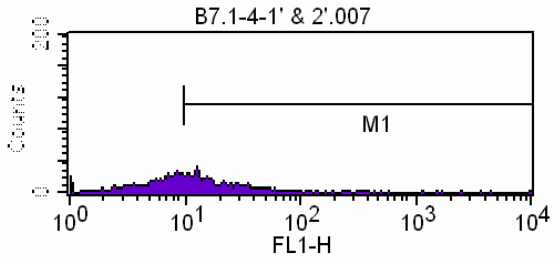


a) b)

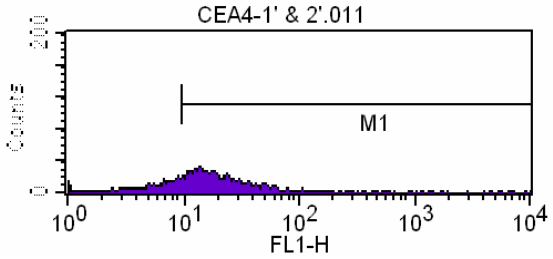
Figure 19. Infection of CT26 at 400 μg/ml G418 on Day 7. a) Neg, CT26 without infection b) CT26 with infection, which became confluent in the 24-well plate.



a)



b)



c)

Figure 20. The CEA antibody in the sera of Balb/c mice. Balb/3T3 was transfected with pAAV-CEA-B7.1 and pAAV-B7.1-IVH3H. The pAAV-CEA-B7.1-IVH3H transfected Balb/3T3 had a higher expression of fluorescence than its negative control and pAAV-B7.1 transfected counterpart by cytometry. The percentage in the M1 region is: a) 1.10% b) 51.17% c) 72.92%, respectively.

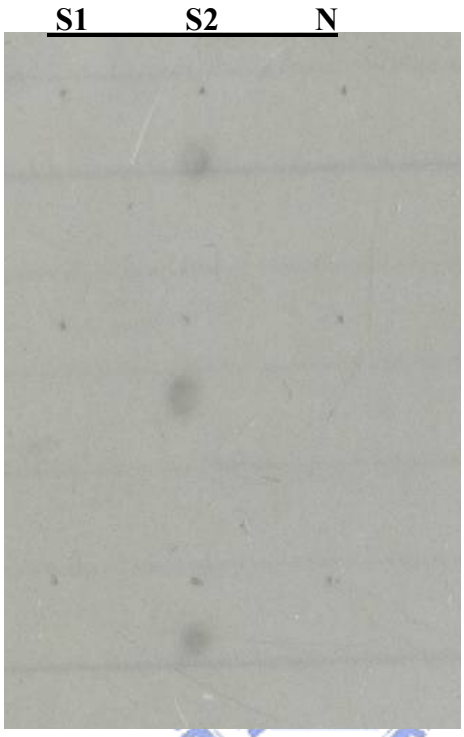


Figure 21. The secretion of CEA from CT26. CEA is secreted by CT26 by dot blotting. N: negative control (PBS alone); S1: cell lysis; S2: cell soup collected after 48 hr incubation at 37℃.

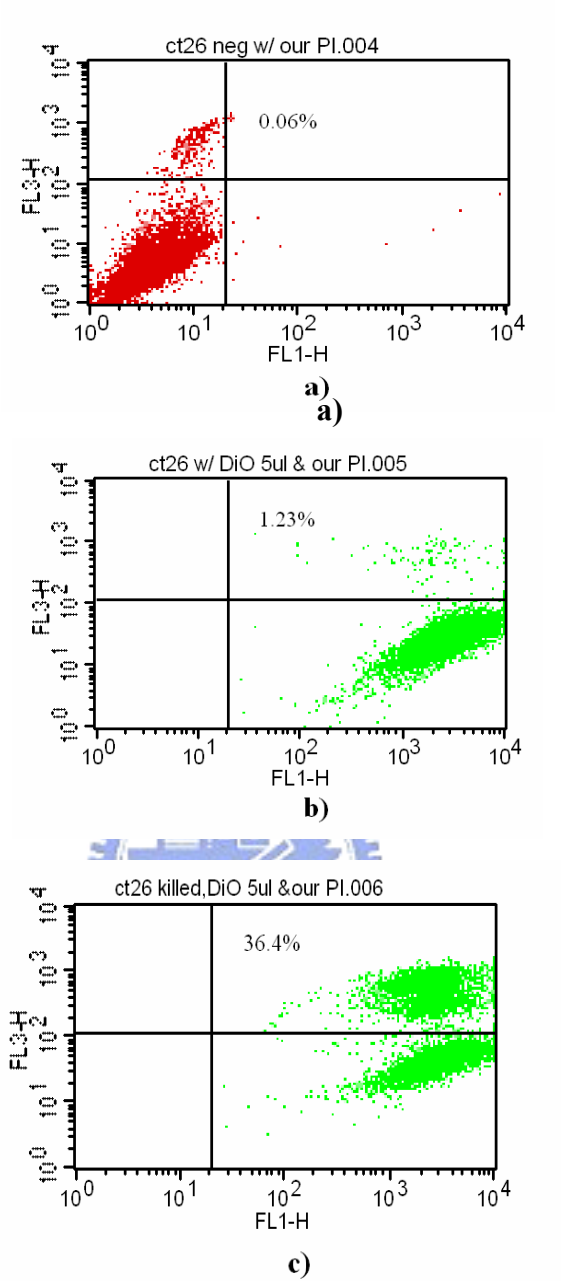
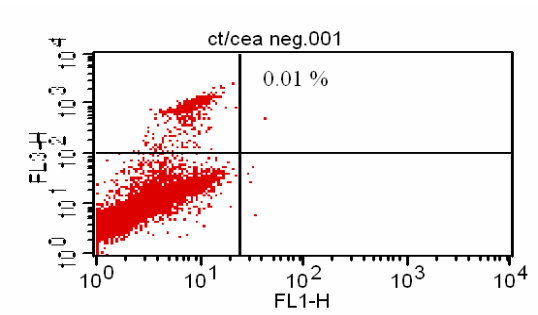
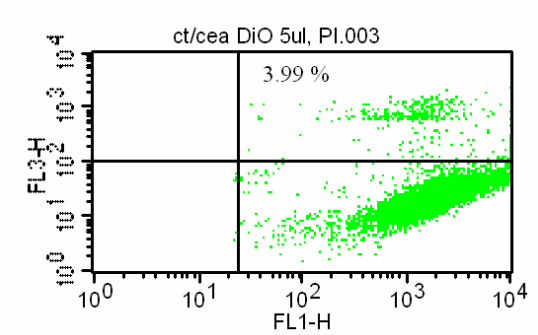


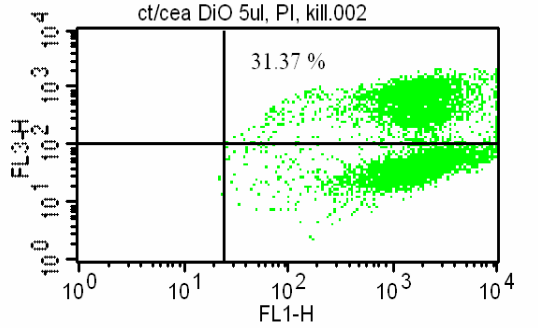
Figure 22. The mortality rate and surface fluorescence of CT26 after overnight DIOC18 staining. Before the killing assay, the effect DIOC18 on CT26 condition was tested. DIOC18 did not affect CT26 after overnight staining as revealed by 50 μg/ml propidium iodide (PI). a) Neg, w/o DIOC18 staining b) DIOC18 staining c) DIOC18 staining + 7% formaldehyde. The mortality rate = (the number of cells of PI staining and DIOC18 staining) / (the number of cells of DIOC18 staining).



a)

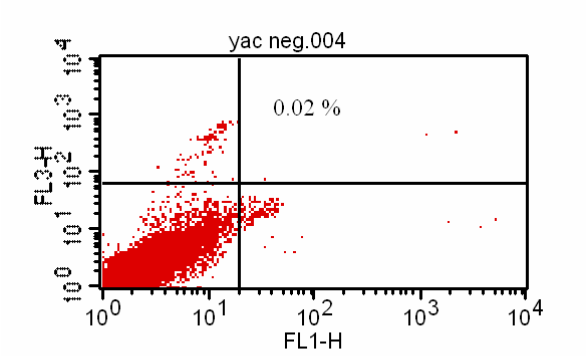


b)

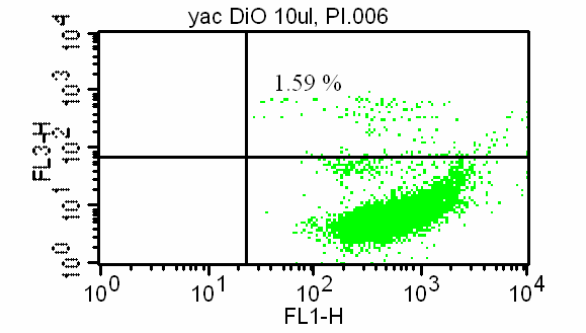


c)

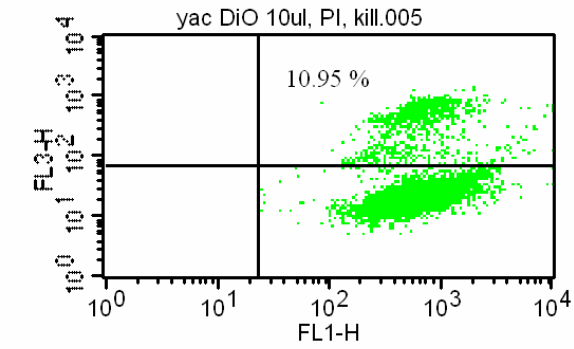
Figure 23. The mortality rate and surface fluorescence of CT26/CEA after overnight DIOC18 staining. Before the killing assay, the effect DIOC18 on CT26/CEA condition was tested. DIOC18 did not affect CT26/CEA after overnight staining as revealed by 50 μg/ml propidium iodide (PI). a) Neg, w/o DIOC18 staining b) DIOC18 staining c) DIOC18 staining + 7% formaldehyde. The mortality rate = (the number of cells of PI staining and DIOC18 staining) / (the number of cells of DIOC18 staining).



a)



b)



c)

Figure 24. The mortality rate and surface fluorescence of YAC-1 after overnight DIOC18 staining. DIOC18 did not affect YAC-1 after overnight staining as revealed by 50 μg/ml propidium iodide (PI). a) Neg, w/o DIOC18 staining b) DIOC18 staining c) DIOC18 staining + 7% formaldehyde. The mortality rate = (the number of cells of PI staining and DIOC18 staining) / (the number of cells of DIOC18 staining).

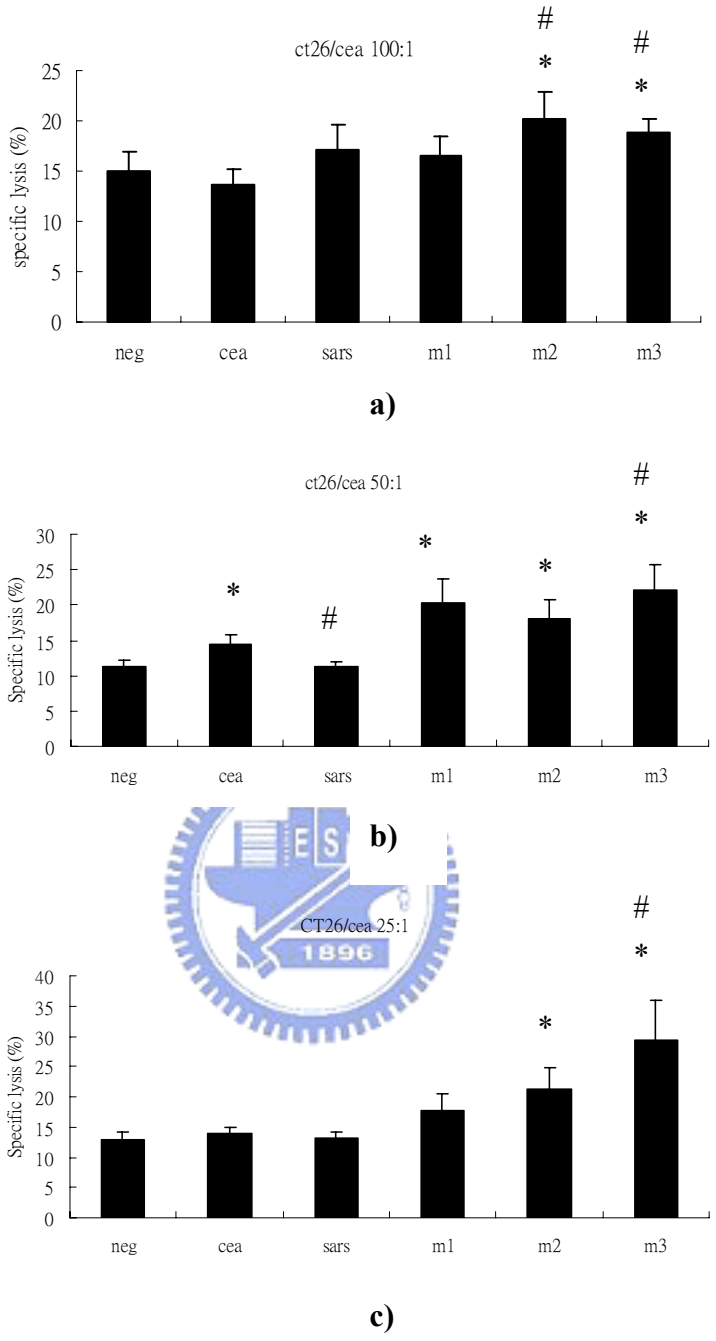


Figure 25. The CT26/CEA killing assay. CT26/CEA was incubated with splenocytes for 4 hr at 37℃. E/T ratios at a) 100/1, b) 50/1, and c) 25/1 are shown. Specific lysis (%) = (the number of cells of PI staining and DIOC18 staining) / (the number of cells of DIOC18 staining). *$p < 0.05$, when compared with the negative group; #$p < 0.05$, when compared with the CEA group.

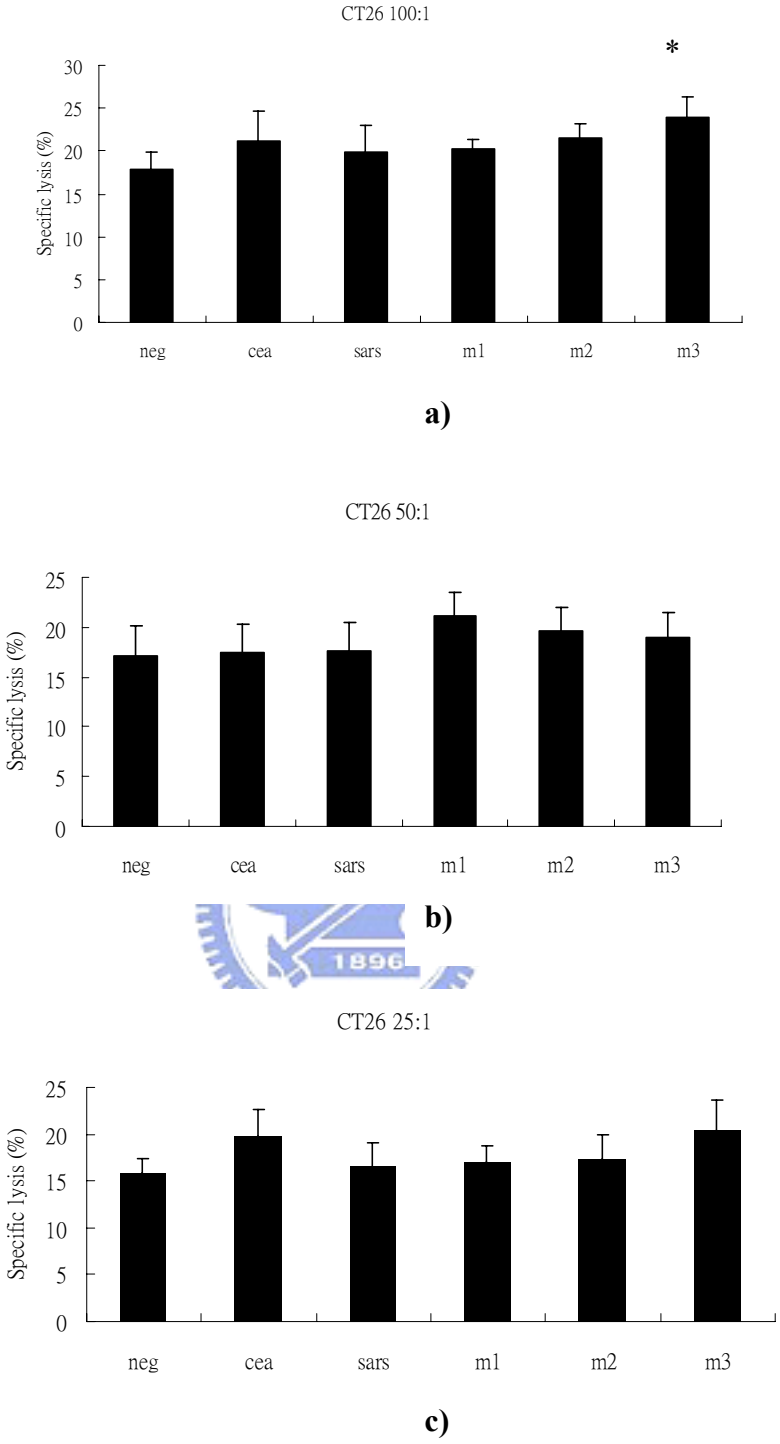


Figure 26. The CT26 killing assay. CT26 was incubated with splenocytes for 4 hr at 37℃. E/T ratios at a) 100/1, b) 50/1, and c) 25/1 are shown. Specific lysis (%) = (the number of cells of PI staining and DIOC18 staining) / (the number of cells of DIOC18 staining). *$p < 0.05$, when compared with the negative group.

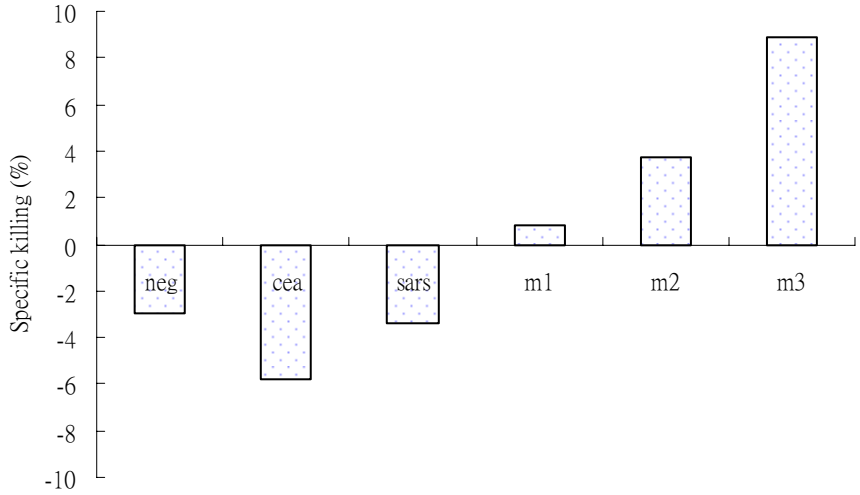


Figure 27. CT26/CEA specific killing at an E/T ratio = 25/1. Specific killing was calculated by the formula below: CT26/CEA specific killing = (% of CT26/CEA killing in **Figure 25c**) – (% of CT26 killing in **Figure 26c**).

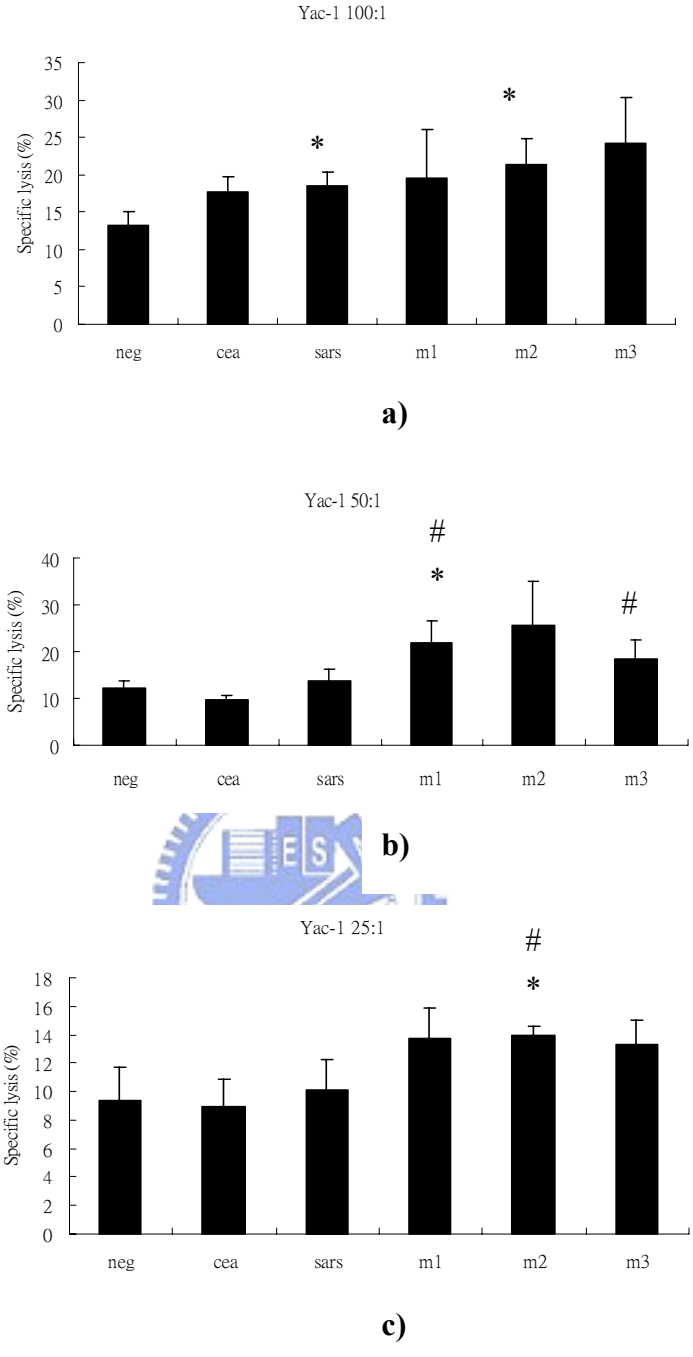
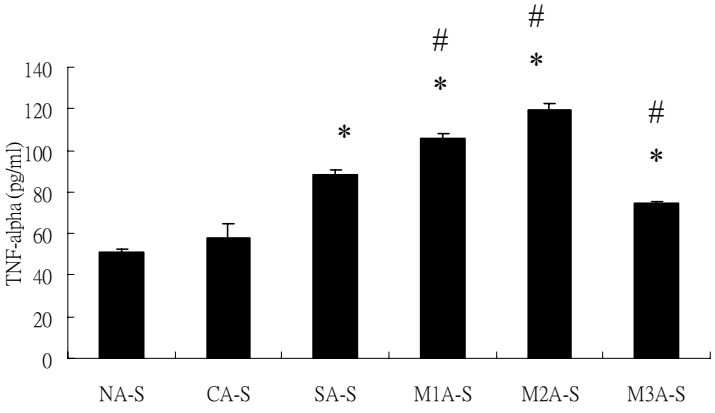
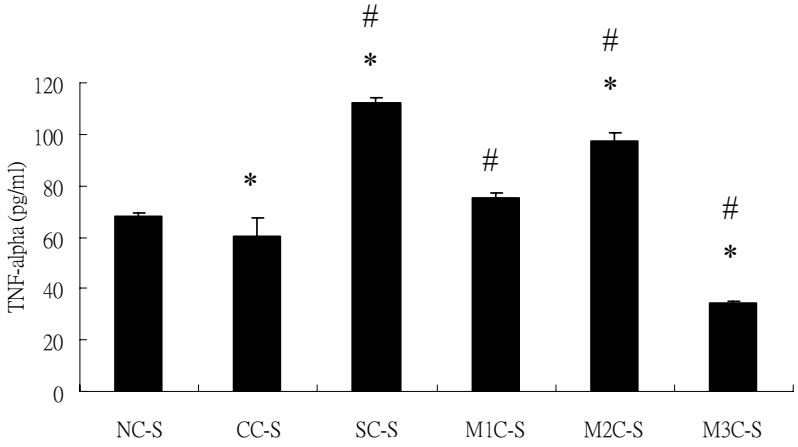


Figure 28. The YAC-1 killing assay. CT26 was incubated with splenocytes for 4 hr at 37℃. E/T ratios at a) 100/1, b) 50/1, and c) 25/1 are shown. Specific lysis (%) = (the number of cells of PI staining and DIOC18 staining) / (the number of cells of DIOC18 staining). *$p < 0.05$, when compared with the negative group; #$p < 0.05$, when compared with the CEA group.



a)



b)

Figure 29. In vitro TNF-α expression after CT26/CEA soup or CT26 soup stimulations. Splenocytes were stimulated by either a) CT26/CEA soup or b) CT26 soup for 24 hr. Groups are shown in the order: negative, CEA, SARS, m1, m2, and m3. *$p < 0.05$, when compared with the negative group; #$p < 0.05$, when compared with the CEA group.

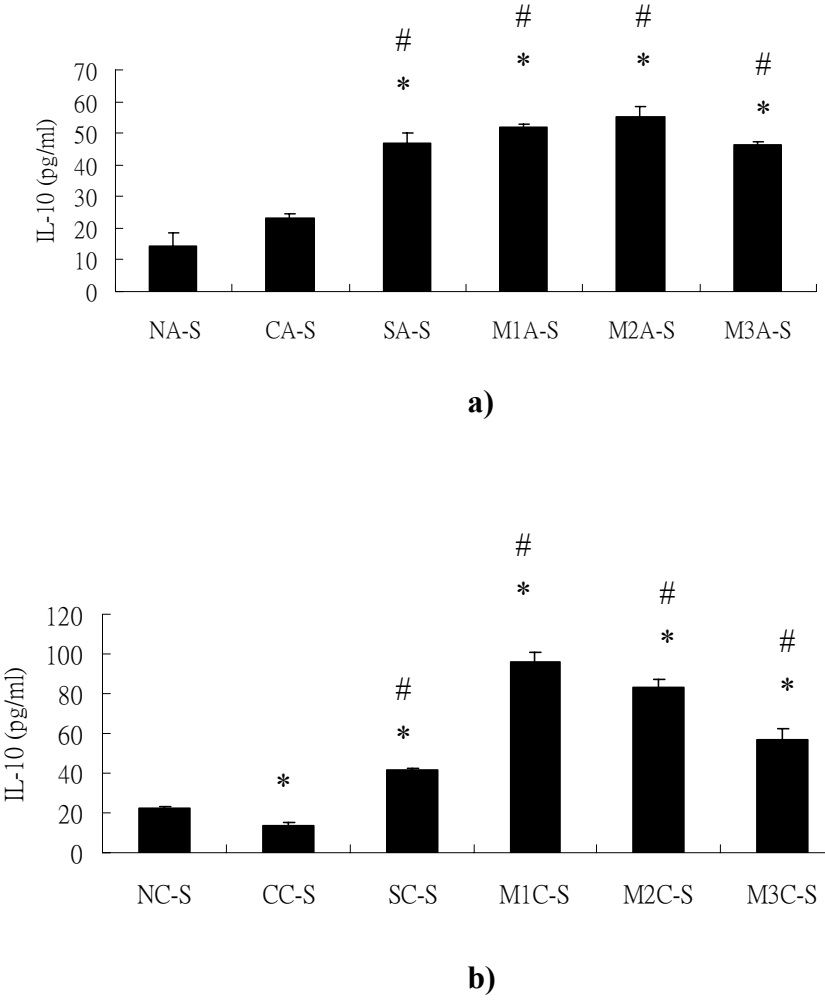


Figure 30. In vitro IL-10 expression after CT26/CEA soup or CT26 soup stimulations. Splenocytes were stimulated by either a) CT26/CEA soup or b) CT26 soup for 24 hr. Groups are shown in the order: negative, CEA, SARS, m1, m2, and m3. *$p < 0.05$, when compared with the negative group; #$p < 0.05$, when compared with the CEA group.

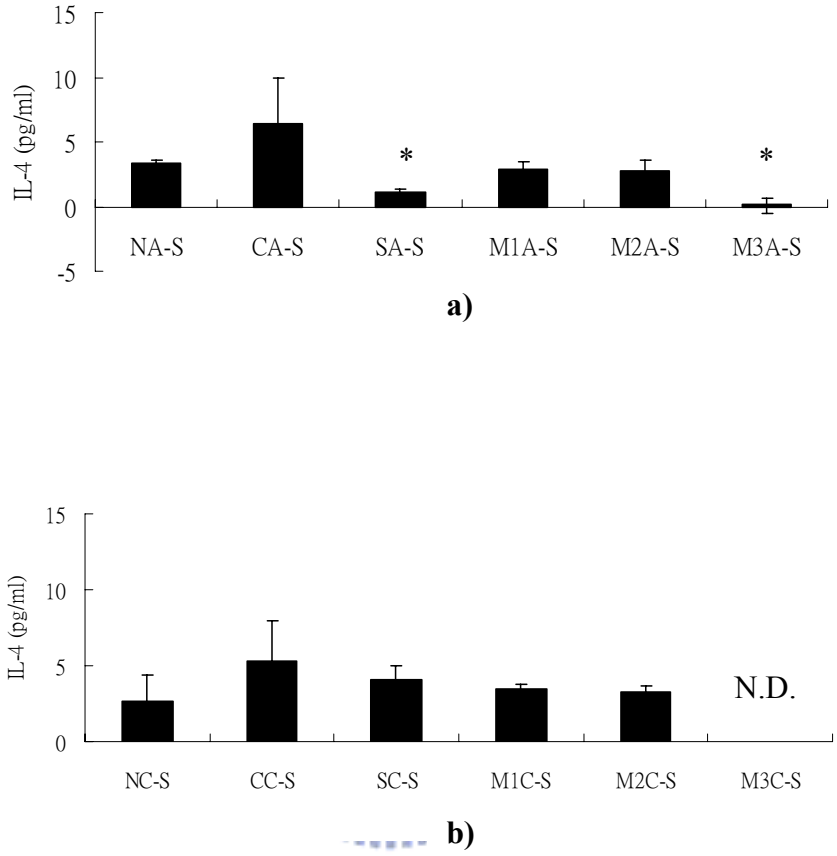


Figure 31. In vitro IL-4 expression after CT26/CEA soup or CT26 soup stimulations. Splenocytes were stimulated by either a) CT26/CEA soup or b) CT26 soup for 24 hr. Groups are shown in the order: negative, CEA, SARS, m1, m2, and m3. *$p < 0.05$, when compared with the negative group; N.D.: non-detectable.

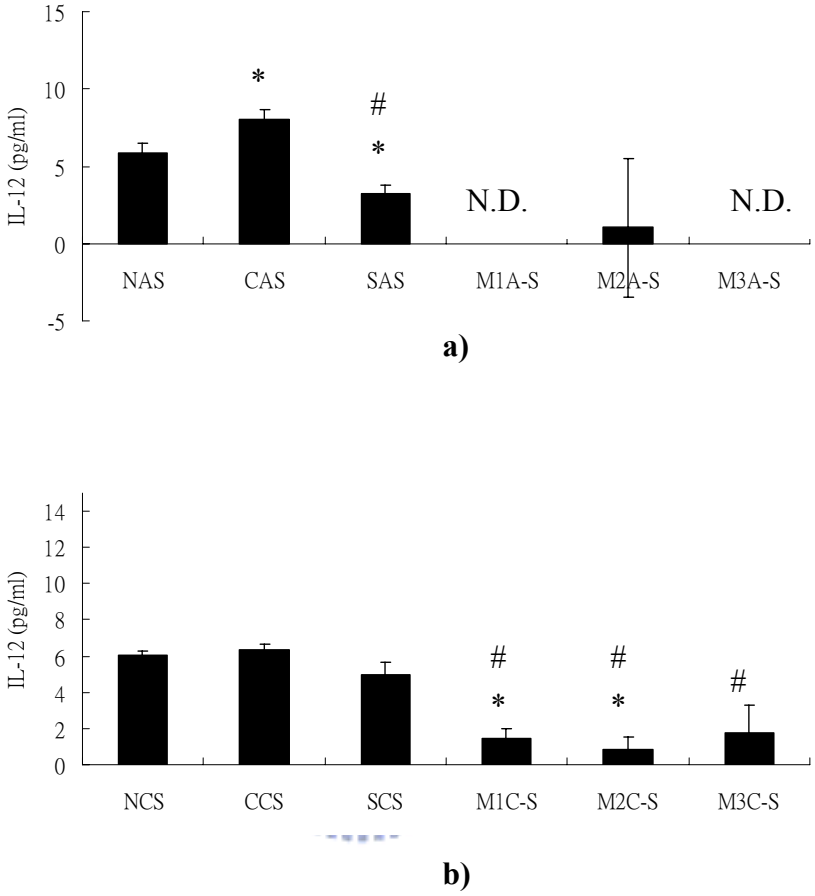


Figure 32. In vitro IL-12 expression after CT26/CEA soup or CT26 soup stimulations. Splenocytes were stimulated by either a) CT26/CEA soup or b) CT26 soup for 24 hr. Groups are shown in the order: negative, CEA, SARS, m1, m2, and m3. *$p < 0.05$, when compared with the negative group; #$p < 0.05$, when compared with the CEA group. ND: non-detectable.

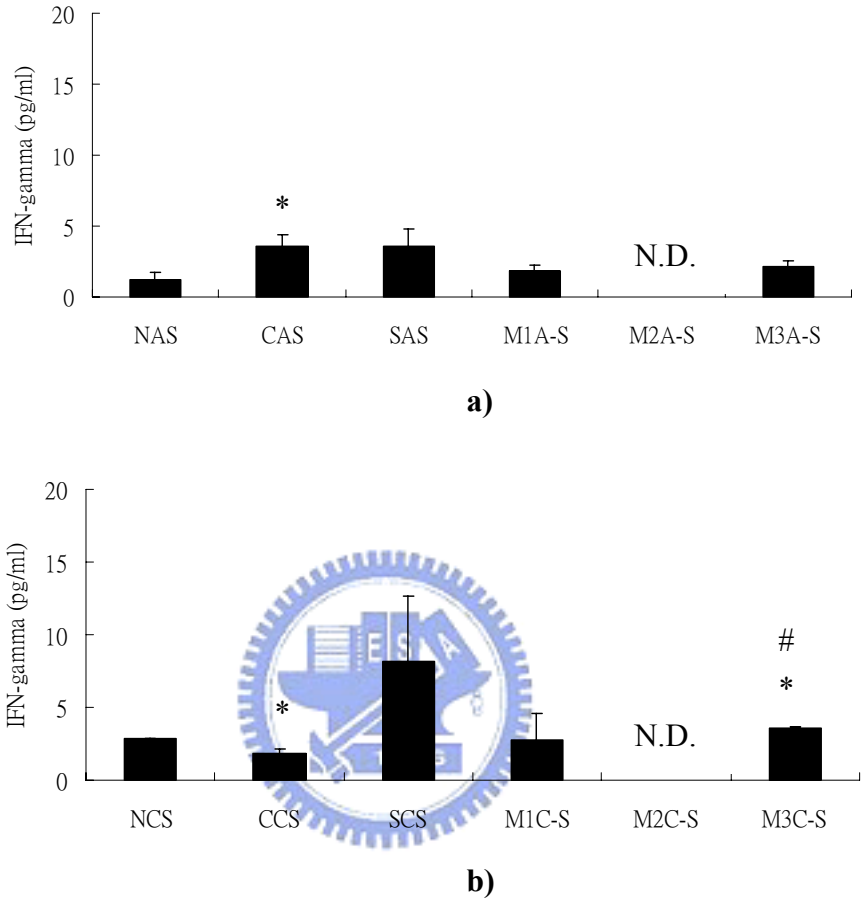


Figure 33. In vitro IFN-γ expression after CT26/CEA soup or CT26 soup stimulations. Splenocytes were stimulated by either a) CT26/CEA soup or b) CT26 soup for 24 hr. Groups are shown in the order: negative, CEA, SARS, m1, m2, and m3. *$p < 0.05$, when compared with the negative group; #$p < 0.05$, when compared with the CEA group. ND: non-detectable.

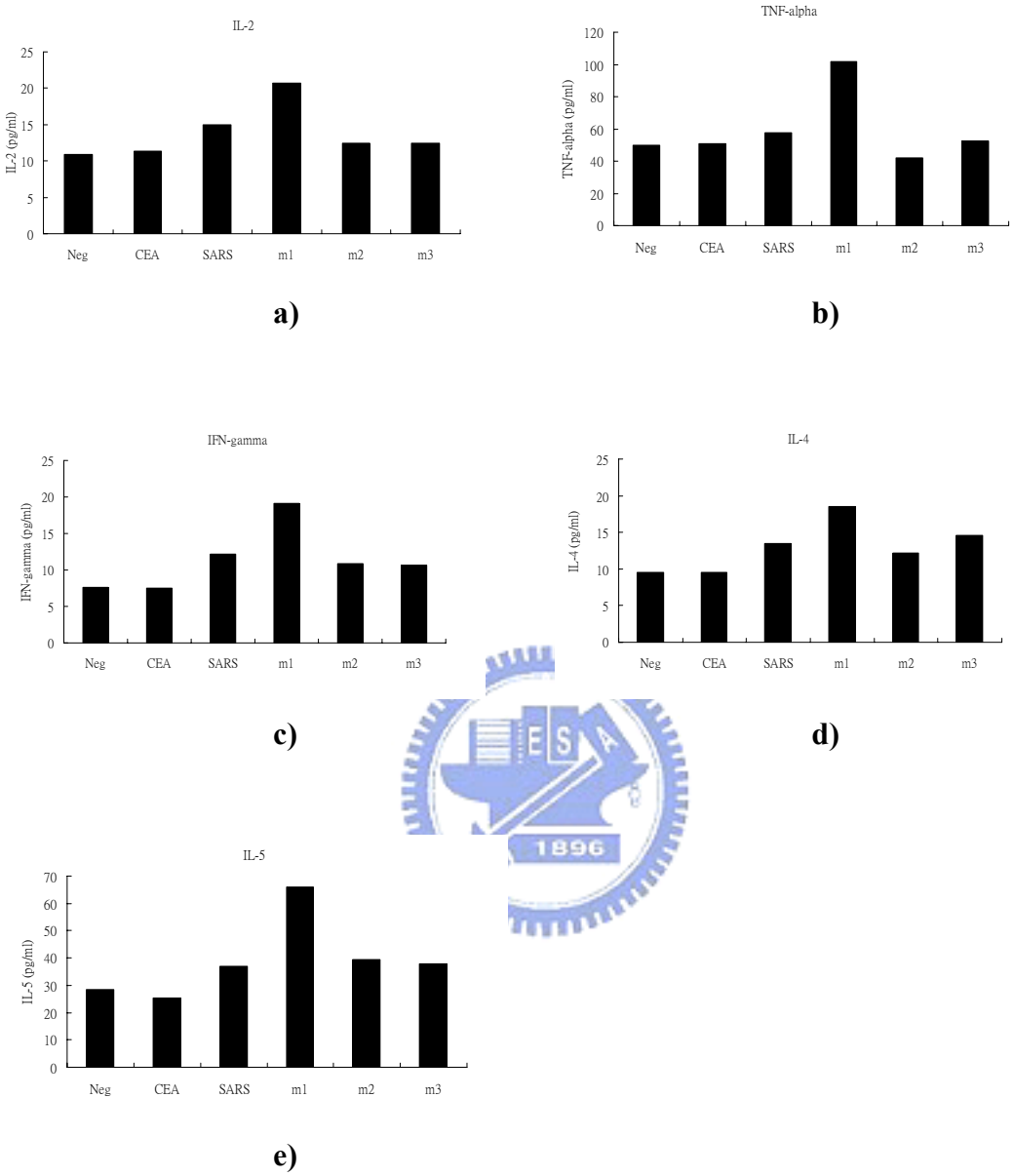


Figure 34. In vivo cytokine expression. Sera were collected for the in vivo cytokine detection, including a) IL-2, b) TNF-α, c) IFN-γ, d) IL-4, and e) IL-5. Groups are shown in the order: negative, CEA, SARS, m1, m2, and m3.

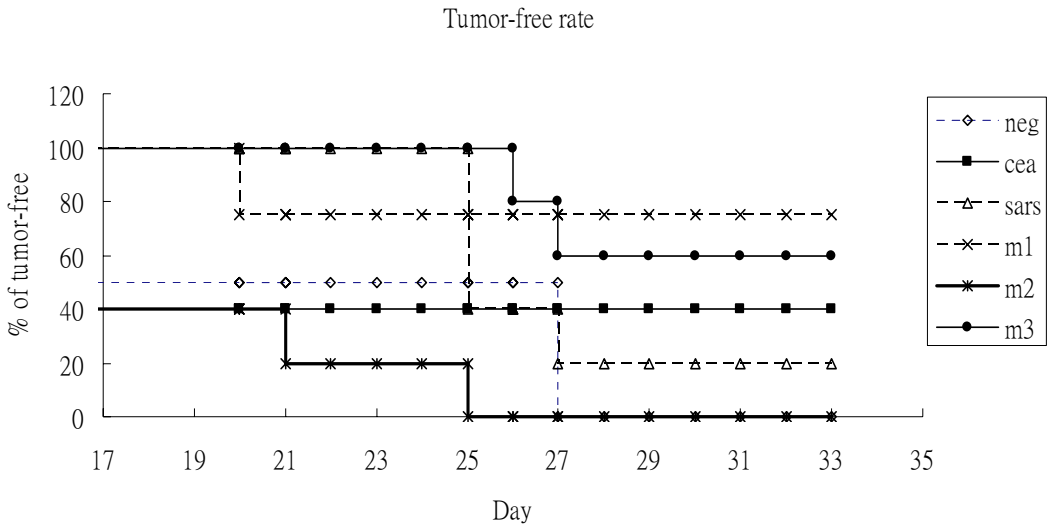


Figure 35. The tumor-free rate of Balb/c mice in the protection assay. Mice were immunized with *Salmonella typhimurium*, which had been transformed into each plasmid construct three times in two weeks and were inoculated with 5 x 10^5 CT26/CEA.

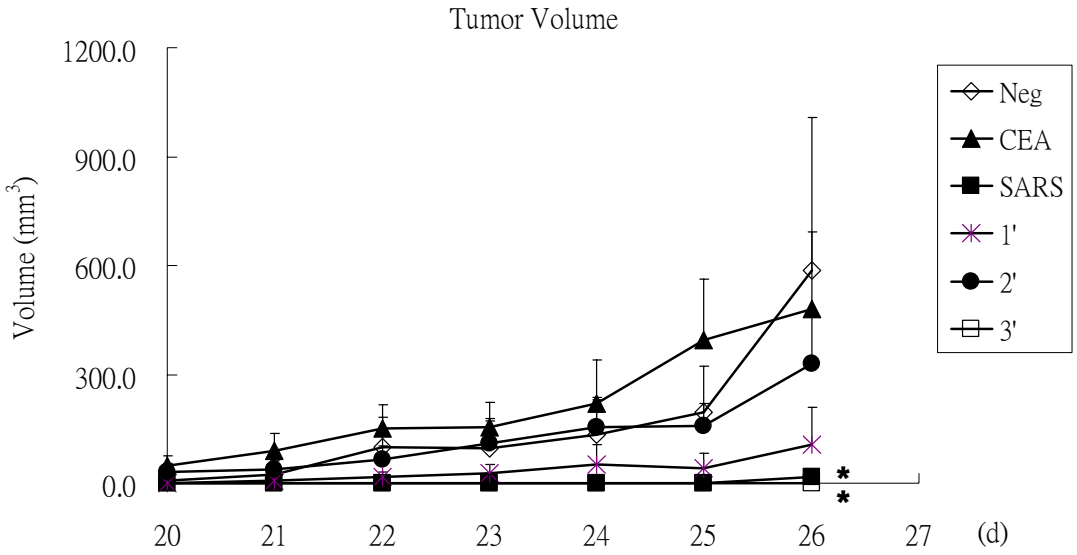


Figure 36. The tumor volume of Balb/c mice in the protection assay. Mice were immunized with *Salmonella typhimurium*, which had been transformed into each plasmid construct three times in two weeks and were inoculated with 5 x 10^5 CT26/CEA. Tumor volume = length x width x height. *$p < 0.05$, when compared with the negative group.

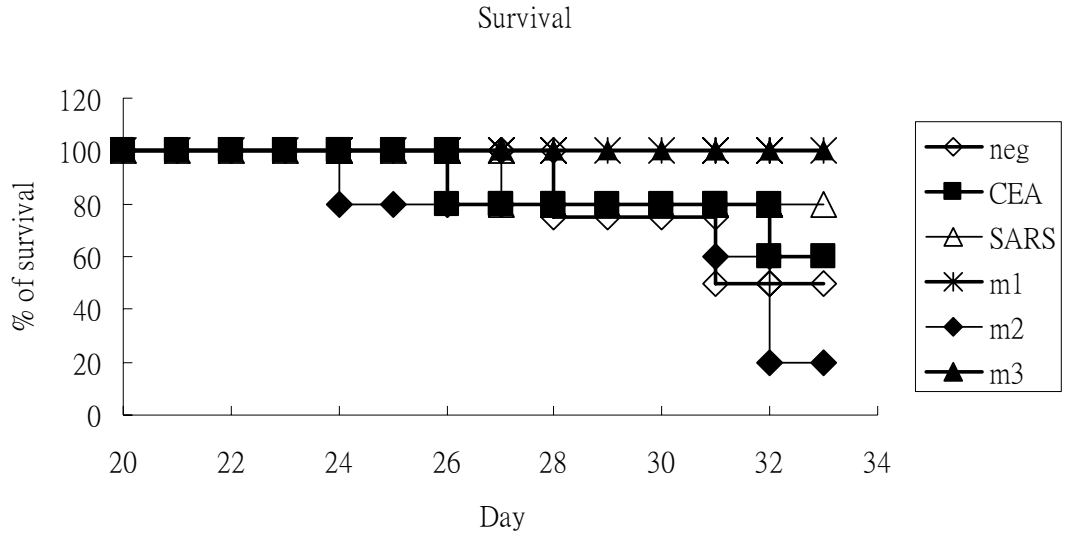


Figure 37. The survival rate of Balb/c mice in the protection assay. Mice were immunized with *Salmonella typhimurium*, which had been transformed into each plasmid construct three times in two weeks and were inoculated with 5 x 10^5 CT26/CEA.

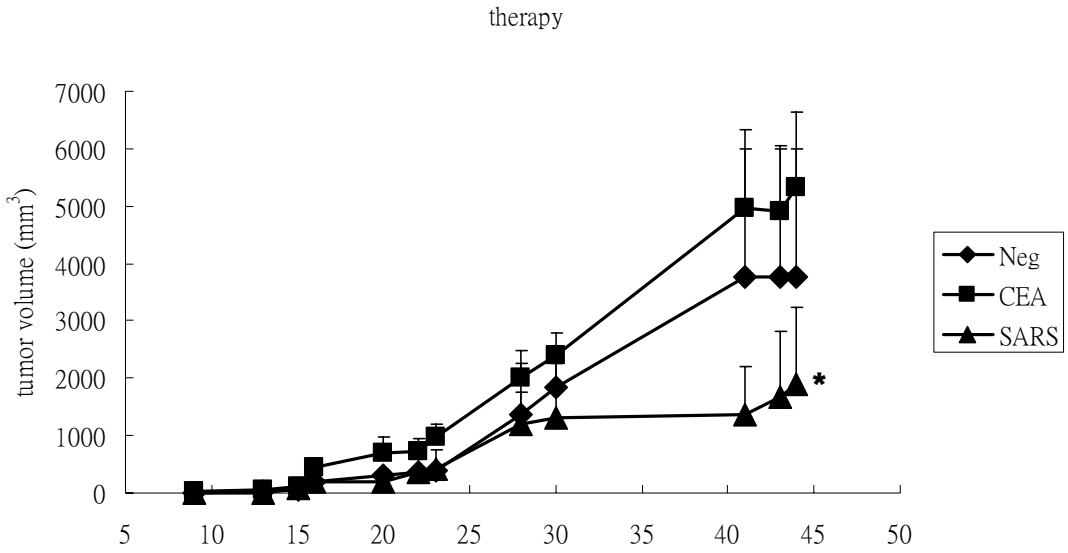


Figure 38. The tumor volume of Balb/c mice in the therapy assay. Mice were inoculated with 1 x 10^5 CT26/CEA and were immunized 4 days later with *Salmonella typhimurium,* which had been transformed into each plasmid construct. Mice were re-boosted once every week and the tumor volume was measured every two to three days. *$p < 0.05$, when compared with the negative group.

Chapter 4 Discussion

This study tested whether the exogenous antigen, the SARS fragment, could induce and enhance the vaccine efficacy of a low immunogenous TAA construct, the CEA fragment, in our animal model. The immunotherapy was realized by orally immunizing Balb/c mice with transformed *Salmonella typhimurium*. Cytokine profile and tumor volume were monitored to understand the effect of the DNA vaccine. Moreover, mutations were introduced on the SARS fragment to enhance the affinity between the MHC molecule and the epitope. Serial accumulation of mutations was designed to compare the protective effect of each construct (m1, m2, m3) with the original fragment (SARS). Mutation affinity was calculated by the Internet software, SYFPEITHI. Unlike other immunotherapies which put an emphasis on TAAs alone, we have set up a platform where a universal antigen is expressed to enhance the immunogenicity of the TAA construct. The concept of utilizing mutations on TAAs for better immunogenicity is applied to the universal antigen. By doing so, we intended to power the strength of the SARS fragment without mutating individual TAA.

The Internet software, SYFPEITHI, has been widely utilized as an epitope prediction algorithm. Pavlenko (Pavlenko, Leder et al. 2005) has proved that epitope specificity of the CTLs was determined by their reactivity against a panel of C-terminus truncated or mutated PSA proteins and use of bioinformatical prediction with the SYFPEITHI algorithm. In the study of Neumann et al. (Neumann, Wagner et al. 2005), the SYFPEITHI algorithm was used to select peptides with a high binding affinity to major histocompatibility complex class 2 (MHC 2) molecules. The pentadecamer epitope p635-649 induced specific CD4+ T-cell responses that were shown to be restricted by HLA-DRB1*1401. The responses could be blocked by preincubation of T cells with anti-CD4 and antigen-presenting cells with anti-HLA-DR, respectively, proving the HLA-DR-restricted presentation of p635-649 and a

CD4+ T-cell-mediated effector response.

The credibility of SYFPEITHI can be justified by the comparison with other software. For example, the transmembrane protein, HM1.24, expressed on the terminal differentiated B cells was scanned for immunogenic peptides using the HLAbinding prediction software SYFPEITHI and BIMAS. Of eight nona-peptides with the highest probability of binding to HLA-A2, the HM1.24 aa22-30 peptide (LLLGIGILV) showed the most frequent activation of CD8+ T cells in healthy volunteers. Moreover, antigen recognition by the HM1.24 aa22-30-specific CD8+ T cells was HLA-A22restricted (Hundemer, Schmidt et al. 2006). Other researches have also utilized the software, SYFPEITHI, BIMAS, or Rankpep, to predict the possible epitopes that can elicit the immune responses (Gomez-Nunez, Pinilla-Ibarz et al. 2006; Molinier-Frenkel, Popa et al. 2006).

We have compared the predicted result with a docking software released by BioXJEM (Yang 2004; Yang 2005). Docking reveals that the sequence of WYVW within the SARS epitope (WYVWLGFIA) predicted by SYFPEITHI makes a large steric hindrance for binding. However, the modifications of the tail part, LGFIA to LGTII (in m2 and m3), can form a better interaction with the MHC molecule at docking, which may partially explain why the CT26/CEA killing is more conspicuous in the m2 and the m3 groups.

To understand the different effect of the DNA vaccine quantitatively, 2 x 10^6 /well splenocytes in each group were stimulated by either CT26/CEA soup (specific stimulation) or CT26 (non-specific stimulation) soup for 24 hr. TNF-α expression was not up-regulated in the CEA group compared with the negative group after CT26/CEA or CT26 soup stimulation. However, the SARS-containing groups had a striking increase in TNF-α expression when compared with the negative and the CEA groups. TNF-α is a proinflammatory and Th1 cytokine. As indicated by Austin et al., TNF-α, IFN-γ, and IL-2 can define cytotoxic T lymphocytes and Th1 effector populations (Austin, Ozawa et al. 1999). Yet, IFN-γ was not significantly increased and IL-2 was undetectable in our experiment.

IL-4, a typical Th2 cytokine, was enhanced in the CEA group compared to the negative group. Yet, it was decreased in the groups with an additional SARS fragment, whether mutated or not. IL-10, however, was raised in the SARS, m1, m2, and m3 groups but not in the CEA group. Though most often it is known as a Th2 cytokine, IL-10 is actually a pleiotropic cytokine with anti-inflammatory, immunosuppressive, immunostimulatory properties (Moore, de Waal Malefyt et al. 2001). It exerts immunostimulatory effects on B cells, cytotoxic T cell development and thymocytes (Conti, Kempuraj et al. 2003). As indicated by Wogensen et al., expression of an IL-10 transgene by insulin-producing pancreatic cells led to an accelerated onset of diabetes in NOD mice (Wogensen, Huang et al. 1993; Wogensen, Lee et al. 1994), with no inhibition of immune-mediated destruction of islets (Lee, Wogensen et al. 1994). In the study of Moritani et al., NOD mice expressing an IL-10 transgene in glucagon-producing pancreatic cells also developed accelerated diabetes (Moritani, Yoshimoto et al. 1994). Consequently, the expression of IL-10 in the model may exert not only a Th2 effect but a Th1 effect as well.

IL-12, though a Th1 cytokine, was slightly up-regulated in the CEA group but not in the SARS group when compared with the negative control after CT26/CEA soup stimulation. It was even more down-regulated in the mutation groups. IL-12 is mostly secreted by macrophages and DCs. The possible mechanism may lie in the fact that the increase in IL-10 in the SARS, m1, m2, and m3 groups exerts an inhibitory effect on macrophages so as to inhibit Th1 activation by blocking macrophage IL-12 synthesis. In the study of Lang et al., an overproduction of inflammatory cytokines and development of chronic inflammatory diseases have been shown in IL-10 gene-deficient mice (Lang, Rutschman et al. 2002).

In general, the CEA antigen only induces a Th2 response whereas the addition of a SARS fragment helps induce and enhance both Th1 and Th2 responses. Th1 is more famous for its antitumor effect. For instance, IL-27 has been proved to possess antiangiogenic and antitumor activities in the B16F10 model. The poorly immunogenic murine melanoma

B16F10 tumors were engineered to overexpress single-chain IL-27 (B16F10 + IL-27). B16F10 + IL-27 cells exerted antitumor activity against not only s.c. tumor but also experimental pulmonary metastasis (Shimizu, Shimamura et al. 2006). On the other hand, the Th2 response has been reported in autologous tumor, where T cells from patients with indolent non-Hodgkin lymphomas frequently showed an activated but apoptosis-prone phenotype (Anichini, Mortarini et al. 2006). It may seem that the Th2 response may down regulate the effect of a Th1 response or contribute less to the antitumor activity, but it has also been observed in some models where a Th2-dominated antitumor immunity has occurred. As reported by Chu Y et al. (Chu, Xia et al. 2006), their DNA vaccine which comprised a modified core peptide of mucin1 (PDTRP) and GM-CSF coding sequence at the C-terminus induced better protection against tumor challenge. The protection is correlated with the type 2 immune responses manifested by an increased IgG1 to IgG2a antibody ratio and a greater induction of GATA-3 and IL-4 mRNA than that of T-bet and IFN-gamma mRNA in spleen cells from vaccinated mice.

Synergistic antitumor effect of both Th1 and Th2 cytokines has also been detected. As indicated by Lopez et al. (Lopez, Adris et al. 2005), the combination of autologous inactivated tumor cells expressing IL-12 and IL-10 induced tumor remission in 50-70% of mice harboring large established colon or mammary tumors and spontaneous lung metastases, with the consequent establishment of an antitumor immune memory. The production of IFN-gamma and IL-4 by spleen cells and the development of tumor-specific IgG1 and IgG2a Abs indicate that each cytokine stimulated its own Th pathway and that both arms were actively engaged in the antitumor effect. The study of IL-21 and IL-15 by Nakano et al. (Nakano, Kishida et al. 2006) provides further evidence for the cellular and humoral responses to tumor cells. IL-21 induced significant elevation of head and neck squamous cell carcinoma-specific CTL activity, while IL-21 and IL-15 augmented NK activity in an additive manner. IL-21 gene transfer also promoted the production of tumor-specific IgG..

Not only does the in vitro cytokine assay reveal an enhanced Th1 and Th2 responses in the SARS-containing groups, but also the in vivo cytokine assay shows the activation of Th1 and Th2 cytokines. On the contrary, CEA alone could not stimulate immune cells to secrete cytokines in vivo. The in vivo assay reveals the total amount of cytokines within an animal, which is the outcome of immunization, suggesting the mechanism of the protection of our DNA vaccine.

The enhancement of Th1 and Th2 in the SARS-containing groups indeed displayed a better protective effect on the animal models. Tumor volume was monitored for 26 days. When compared with the negative group, mice immunized with CEA alone did not show any significant difference in the protection assay and in the therapy assay. However, the SARS and m3 groups had the smallest tumor volume in the protection assay. The m1 construct also provided enough protection for mice, but the m2 construct did not seem to be able to sufficiently suppress tumor growth.

It was expected that the m2 construct should have a higher antitumor activity than the m1 construct. However, the protection assay reveals that it did not. Some reasons are given to explain the phenomenon. First of all, as introduced in 1.3.3, beside the binding affinity between the epitope and the MHC molecule, there exists another binding affinity between the epitope-MHC complex and the TCR. Some researches have demonstrated that such an affinity may influence the activation of cytotoxic T cells as well (Slansky, Rattis et al. 2000). Second, only nonamers were chosen as the epitopes in the software calculation. But other sequence length may also be generated in the animal. It is not known if such a nonamer calculation is the best outcome in vivo. Furthermore, there are other MHC alleles to accommodate a variety of epitopes in a mouse. For example, besides the H2-K form, H2-D and H2-L belong to the MHC class I molecules as well. Therefore, it is beyond our control to generate an epitope mutation that induces the best immunity in vivo. Some other unknown mechanisms may undermine the protective effect of such an immune strategy since immunity itself is interlaced

to reach homeostasis.

The truncated CEA antigen successfully simulates a low immunogenous TAA, which is often the case in most tumor cells. It did not induce any favorable cytokines in vitro and in vivo to enhance either Th1 or Th2 responses. Its killing assay reveals that CEA alone could not specifically kill CT26/CEA and its innate immunity (YAC-1 killing) did not show any significant difference when compared with the negative control. However, SARS fusion not only induces but also enhances anti-CEA activity, as shown in the CT26/CEA specific killing. Moreover, specific killing is proportional to the number of mutations, indicating that our mutations did work well in vitro. The antitumor activity of the m2 construct is limited in vivo due to the reasons provided above. But all in all, the foreign parental or mutated SARS fragments could enhance the anti-tumor efficacy of the tumor vaccine against endogenous tumor antigens. As a result, we provide a platform to enhance the adjuvant effect of the foreign peptide by computer prediction.

References

Anichini, A., R. Mortarini, et al. (2006). "Skewed T-cell differentiation in patients with indolent non-Hodgkin lymphoma reversed by ex vivo T-cell culture with gammac cytokines." Blood **107**(2): 602-9.

Austin, L. M., M. Ozawa, et al. (1999). "The majority of epidermal T cells in Psoriasis vulgaris lesions can produce type 1 cytokines, interferon-gamma, interleukin-2, and tumor necrosis factor-alpha, defining TC1 (cytotoxic T lymphocyte) and TH1 effector populations: a type 1 differentiation bias is also measured in circulating blood T cells in psoriatic patients." J Invest Dermatol **113**(5): 752-9.

Bakker, A. B., S. H. van der Burg, et al. (1997). "Analogues of CTL epitopes with improved MHC class-I binding capacity elicit anti-melanoma CTL recognizing the wild-type epitope." Int J Cancer **70**(3): 302-9.

Banchereau, J. and R. M. Steinman (1998). "Dendritic cells and the control of immunity." Nature **392**(6673): 245-52.

Ben-Neriah, Y., G. Q. Daley, et al. (1986). "The chronic myelogenous leukemia-specific P210 protein is the product of the bcr/abl hybrid gene." Science **233**(4760): 212-4.

Boczkowski, D., S. K. Nair, et al. (1996). "Dendritic cells pulsed with RNA are potent antigen-presenting cells in vitro and in vivo." J Exp Med **184**(2): 465-72.

Boon, T., J. C. Cerottini, et al. (1994). "Tumor antigens recognized by T lymphocytes." Annu Rev Immunol **12**: 337-65.

Bueler, H. and R. C. Mulligan (1996). "Induction of antigen-specific tumor immunity by genetic and cellular vaccines against MAGE: enhanced tumor protection by coexpression of granulocyte-macrophage colony-stimulating factor and B7-1." Mol Med **2**(5): 545-55.

Chen, C. H. and T. C. Wu (1998). "Experimental vaccine strategies for cancer immunotherapy." J Biomed Sci **5**(4): 231-52.

Chen, Y., D. Hu, et al. (1998). "DNA vaccines encoding full-length or truncated Neu induce protective immunity against Neu-expressing mammary tumors." Cancer Res **58**(9): 1965-71.

Chu, Y., M. Xia, et al. (2006). "Th2-dominated antitumor immunity induced by DNA immunization with the genes coding for a basal core peptide PDTRP and GM-CSF." Cancer Gene Ther **13**(5): 510-9.

Clairmont, C., K. C. Lee, et al. (2000). "Biodistribution and genetic stability of the novel antitumor agent VNP20009, a genetically modified strain of Salmonella typhimurium." J Infect Dis **181**(6): 1996-2002.

Condon, C., S. C. Watkins, et al. (1996). "DNA-based immunization by in vivo transfection of dendritic cells." Nat Med **2**(10): 1122-8.

Conti, P., D. Kempuraj, et al. (2003). "IL-10, an inflammatory/inhibitory cytokine, but not always." Immunol Lett **86**(2): 123-9.

Cox, A. L., J. Skipper, et al. (1994). "Identification of a peptide recognized by five melanoma-specific human cytotoxic T cell lines." Science **264**(5159): 716-9.

Dougan G., H. C., Maskell DJ. (1986). "Live oral Salmonella vaccines: potential use of attenuated strains carriers of heterologous antigen to the immune system." Parasite Immunol(9): 151-160.

Dyall, R., W. B. Bowne, et al. (1998). "Heteroclitic immunization induces tumor immunity." J Exp Med **188**(9): 1553-61.

Fagan, P. K., S. P. Djordjevic, et al. (1997). "Oral immunization of mice with attenuated Salmonella typhimurium aroA expressing a recombinant Mycoplasma hyopneumoniae antigen (NrdF)." Infect Immun **65**(6): 2502-7.

Fahrer, A. M., H. M. Geysen, et al. (1995). "Analysis of the requirements for class II-restricted T cell recognition of a single determinant reveals considerable diversity in the T cell response and degeneracy of peptide binding to I-Ed." J Immunol **155**(6): 2849-57.

Gervois, N., Y. Guilloux, et al. (1996). "Suboptimal activation of melanoma infiltrating lymphocytes (TIL) due to low avidity of TCR/MHC-tumor peptide interactions." J Exp Med **183**(5): 2403-7.

Gillespie, G. Y. and S. W. Russell (1978). "Development and persistence of cytolytic T lymphocytes in regressing or progressing Moloney sarcomas." Int J Cancer **21**(1): 94-99.

Gomez-Nunez, M., J. Pinilla-Ibarz, et al. (2006). "Peptide binding motif predictive algorithms correspond with experimental binding of leukemia vaccine candidate peptides to HLA-A*0201 molecules." Leuk Res.

Gong, J., L. Chen, et al. (1997). "Induction of antigen-specific antitumor immunity with adenovirus-transduced dendritic cells." Gene Ther **4**(10): 1023-8.

Halaban, R., S. H. Pomerantz, et al. (1983). "Regulation of tyrosinase in human melanocytes grown in culture." J Cell Biol **97**(2): 480-8.

Hans-Georg Rammensee, J. B., Niels Nikolaus Emmerich, Oskar Alexander Bachor, Stefan Stevanovic (1999). SYFPEITHI: database for MHC ligands and peptide motifs.

Hormaeche CE, K. C., Mastroeni P, Villarreal B, Dougan G, Roberts M, et al. (1995). Molecular and clinical aspects of bacterial vaccine development, John Wiley & Sons Press.

Huang, K. J., I. J. Su, et al. (2005). "An interferon-gamma-related cytokine storm in SARS patients." J Med Virol **75**(2): 185-94.

Hundemer, M., S. Schmidt, et al. (2006). "Identification of a new HLA-A2-restricted T-cell epitope within HM1.24 as immunotherapy target for multiple myeloma." Exp Hematol **34**(4): 486-96.

Jain, A., J. E. Slansky, et al. (2003). "Synergistic effect of a granulocyte-macrophage colony-stimulating factor-transduced tumor vaccine and systemic interleukin-2 in the treatment of murine colorectal cancer hepatic metastases." Ann Surg Oncol **10**(7): 810-20.

Janeway, t., Walport, Shlomchik (2005). Immunobiology. New York, Garland Science Publishing.

King, R. J. B. (2000). Cancer Biology. Harlow, Pearson Education Limited.

King, R. J. B. (2000). Oncogenes. Cancer Biology. Harlow, Pearson Education Limited**:** 79-80.

King, R. J. B. (2000). Oncogenes, repressor genes and viruses. Cancer Biology. Harlow, Pearson Education Limited**:** 72.

King, R. J. B. (2000). Repressor genes. Cancer Biology. Harlow, Pearson Education Limited**:** 89.

Kops, S. K., et al. (1997). "Salmonella typhimurium as an anti-cancer vector: localization within solid tumors. ." Proc. Am. Assoc. Cancer Res. **38**: 7.

Kounalakis, N. and J. S. Goydos (2005). "Tumor cell and circulating markers in melanoma: diagnosis, prognosis, and management." Curr Oncol Rep **7**(5): 377-82.

Lang, R., R. L. Rutschman, et al. (2002). "Autocrine deactivation of macrophages in transgenic mice constitutively overexpressing IL-10 under control of the human CD68 promoter." J Immunol **168**(7): 3402-11.

Lee, M. S., L. Wogensen, et al. (1994). "Pancreatic islet production of murine interleukin-10 does not inhibit immune-mediated tissue destruction." J Clin Invest **93**(3): 1332-8.

Lee, P. P., C. Yee, et al. (1999). "Characterization of circulating T cells specific for tumor-associated antigens in melanoma patients." Nat Med **5**(6): 677-85.

Li, D., H. Li, et al. (2006). "Heat shock fusion protein induces both specific and nonspecific anti-tumor immunity." Eur J Immunol **36**(5): 1324-36.

Li, Q., Y. Zhang, et al. (2006). "[Immunotherapeutic efficacy of both helper T lymphocytes and cytotoxic T lymphocytes epitopes augmented dendritic cells tumor vaccine on gastric cancer]." Zhonghua Wei Chang Wai Ke Za Zhi **9**(2): 148-51.

Liu, D. L., M. Q. Yang, et al. (1996). "Repeated immunotherapy using intratumoural injection with recombinant interleukin-2 and tumour-infiltrating lymphocytes inhibits growth of breast cancer and induces apoptosis of tumour cells." Cancer Lett **103**(2): 131-6.

Lopez, M. V., S. K. Adris, et al. (2005). "IL-12 and IL-10 expression synergize to induce the immune-mediated eradication of established colon and mammary tumors and lung metastasis." J Immunol **175**(9): 5885-94.

Low K, I. M., Lin S, et al (1999). "VNP2009, a genetically modified Salmonella typhimurium for treatment of solid tumors." Proc Natl Acad Sci U S A **40**: 87.

Low, K. B., M. Ittensohn, et al. (1999). "Lipid A mutant Salmonella with suppressed virulence and TNFalpha induction retain tumor-targeting in vivo." Nat Biotechnol **17**(1): 37-41.

Luo X, I. M., Liu Y, et al. (1999). "Genetically modified Salmonella typhimurium inhibited growth of primary tumors and metastases." Proc Natl Acad Sci U S A **40**: 87 (Abstr.).

Lurquin, C., A. Van Pel, et al. (1989). "Structure of the gene of tum- transplantation antigen P91A: the mutated exon encodes a peptide recognized with Ld by cytolytic T cells." Cell **58**(2): 293-303.

Marshall, D. J., K. A. Rudnick, et al. (2006). "Interleukin-18 enhances Th1 immunity and tumor protection of a DNA vaccine." Vaccine **24**(3): 244-53.

Matsumura, M., M. Ijichi, et al. (2001). "Simple quantitative assay of alpha-fetoprotein mRNA in liver tissue using the real-time detection polymerase chain reaction assay - its application for clinical use." Hepatol Res **20**(1): 84-96.

Mazurek, C. and J. F. Duplan (1959). "[Stimulation & inhibition of the growth of Ehrlich's ascites carcinoma by irradiated tumor cells.]." Bull Assoc Fr Etud Cancer **46**(1): 119-31.

McCoy, M. S., C. I. Bargmann, et al. (1984). "Human colon carcinoma Ki-ras2 oncogene and its corresponding proto-oncogene." Mol Cell Biol **4**(8): 1577-82.

Molinier-Frenkel, V., N. Popa, et al. (2006). "Analysis of the intraclonal diversity of HLA-A0201-restricted T lymphocyte epitopes in follicular lymphoma idiotype." Br J Haematol **132**(4): 459-68.

Moore, K. W., R. de Waal Malefyt, et al. (2001). "Interleukin-10 and the interleukin-10 receptor." Annu Rev Immunol **19**: 683-765.

Morgan, D. J., H. T. Kreuwel, et al. (1998). "Activation of low avidity CTL specific for a self epitope results in tumor rejection but not autoimmunity." J Immunol **160**(2): 643-51.

Moritani, M., K. Yoshimoto, et al. (1994). "Transgenic expression of IL-10 in pancreatic islet A cells accelerates autoimmune insulitis and diabetes in non-obese diabetic mice." Int Immunol **6**(12): 1927-36.

Muraro, S., A. Bondanza, et al. (2005). "Molecular modification of idiotypes from B-cell lymphomas for expression in mature dendritic cells as a strategy to induce tumor-reactive CD4+ and CD8+ T-cell responses." Blood **105**(9): 3596-604.

Nakano, H., T. Kishida, et al. (2006). "Interleukin-21 triggers both cellular and humoral immune responses leading to therapeutic antitumor effects against head and neck squamous cell carcinoma." J Gene Med **8**(1): 90-9.

Neumann, F., C. Wagner, et al. (2005). "Identification of an epitope derived from the cancer testis antigen HOM-TES-14/SCP1 and presented by dendritic cells to circulating CD4+ T cells." Blood **106**(9): 3105-13.

Pardoll, D. M. (1994). "Tumour antigens. A new look for the 1990s." Nature **369**(6479): 357.

Parkhurst, M. R., M. L. Salgaller, et al. (1996). "Improved induction of melanoma-reactive CTL with peptides from the melanoma antigen gp100 modified at HLA-A*0201-binding residues." J Immunol **157**(6): 2539-48.

Pavlenko, M., C. Leder, et al. (2005). "Identification of an immunodominant

H-2D(b)-restricted CTL epitope of human PSA." Prostate **64**(1): 50-9.

Pawelek, J. M., K. B. Low, et al. (1997). "Tumor-targeted Salmonella as a novel anticancer vector." Cancer Res **57**(20): 4537-44.

Ribas, A., L. H. Butterfield, et al. (1997). "Genetic immunization for the melanoma antigen MART-1/Melan-A using recombinant adenovirus-transduced murine dendritic cells." Cancer Res **57**(14): 2865-9.

Rosenberg, S. A. (1996). "The immunotherapy of solid cancers based on cloning the genes encoding tumor-rejection antigens." Annu Rev Med **47**: 481-91.

Rosenberg, S. A., J. C. Yang, et al. (1998). "Immunologic and therapeutic evaluation of a synthetic peptide vaccine for the treatment of patients with metastatic melanoma." Nat Med **4**(3): 321-7.

Salopek, T. G. and K. Jimbow (1996). "Induction of melanogenesis during the various melanoma growth phases and the role of tyrosinase, lysosome-associated membrane proteins, and p90 calnexin in the melanogenesis cascade." J Investig Dermatol Symp Proc **1**(2): 195-202.

Scholl, S. M., J. M. Balloul, et al. (2000). "Recombinant vaccinia virus encoding human MUC1 and IL2 as immunotherapy in patients with breast cancer." J Immunother **23**(5): 570-80.

Schultz, J., J. Pavlovic, et al. (2000). "Immune modulation in cancer using DNA inoculation--antitumour effect of interleukin-12." Dev Biol (Basel) **104**: 109-14.

Shibata, H. R., L. M. Jerry, et al. (1976). "Immunotherapy of human malignant melanoma with irradiated tumor cells, oral Bacillus Calmette-Guerin, and levamisole." Ann N Y Acad Sci **277**(00): 355-66.

Shimizu, M., M. Shimamura, et al. (2006). "Antiangiogenic and antitumor activities of IL-27." J Immunol **176**(12): 7317-24.

Slansky, J. E., F. M. Rattis, et al. (2000). "Enhanced antigen-specific antitumor immunity with altered peptide ligands that stabilize the MHC-peptide-TCR complex." Immunity **13**(4): 529-38.

Somerville, J. E., Jr., L. Cassiano, et al. (1999). "Escherichia coli msbB gene as a virulence factor and a therapeutic target." Infect Immun **67**(12): 6583-90.

Song, E. S., V. Lee, et al. (1997). "Antigen presentation in retroviral vector-mediated gene transfer in vivo." Proc Natl Acad Sci U S A **94**(5): 1943-8.

Song, W., H. L. Kong, et al. (1997). "Dendritic cells genetically modified with an adenovirus vector encoding the cDNA for a model antigen induce protective and therapeutic antitumor immunity." J Exp Med **186**(8): 1247-56.

Specht, J. M., G. Wang, et al. (1997). "Dendritic cells retrovirally transduced with a model antigen gene are therapeutically effective against established pulmonary metastases." J Exp Med **186**(8): 1213-21.

Steinman, R. M. (1991). "The dendritic cell system and its role in immunogenicity." Annu

Rev Immunol **9**: 271-96.

Sznol, M., S. L. Lin, et al. (2000). "Use of preferentially replicating bacteria for the treatment of cancer." J Clin Invest **105**(8): 1027-30.

Takahashi, T., M. M. Nau, et al. (1989). "p53: a frequent target for genetic abnormalities in lung cancer." Science **246**(4929): 491-4.

Thompson, J. A., F. Grunert, et al. (1991). "Carcinoembryonic antigen gene family: molecular biology and clinical perspectives." J Clin Lab Anal **5**(5): 344-66.

Timmerman, J. M. and R. Levy (1999). "Dendritic cell vaccines for cancer immunotherapy." Annu Rev Med **50**: 507-29.

Waeckerle-Men, Y., E. Uetz-von Allmen, et al. (2006). "Dendritic cell-based multi-epitope immunotherapy of hormone-refractory prostate carcinoma." Cancer Immunol Immunother.

Wang, R. F. and S. A. Rosenberg (1996). "Human tumor antigens recognized by T lymphocytes: implications for cancer therapy." J Leukoc Biol **60**(3): 296-309.

Wogensen, L., X. Huang, et al. (1993). "Leukocyte extravasation into the pancreatic tissue in transgenic mice expressing interleukin 10 in the islets of Langerhans." J Exp Med **178**(1): 175-85.

Wogensen, L., M. S. Lee, et al. (1994). "Production of interleukin 10 by islet cells accelerates immune-mediated destruction of beta cells in nonobese diabetic mice." J Exp Med **179**(4): 1379-84.

Wong, C. K., C. W. Lam, et al. (2004). "Plasma inflammatory cytokines and chemokines in severe acute respiratory syndrome." Clin Exp Immunol **136**(1): 95-103.

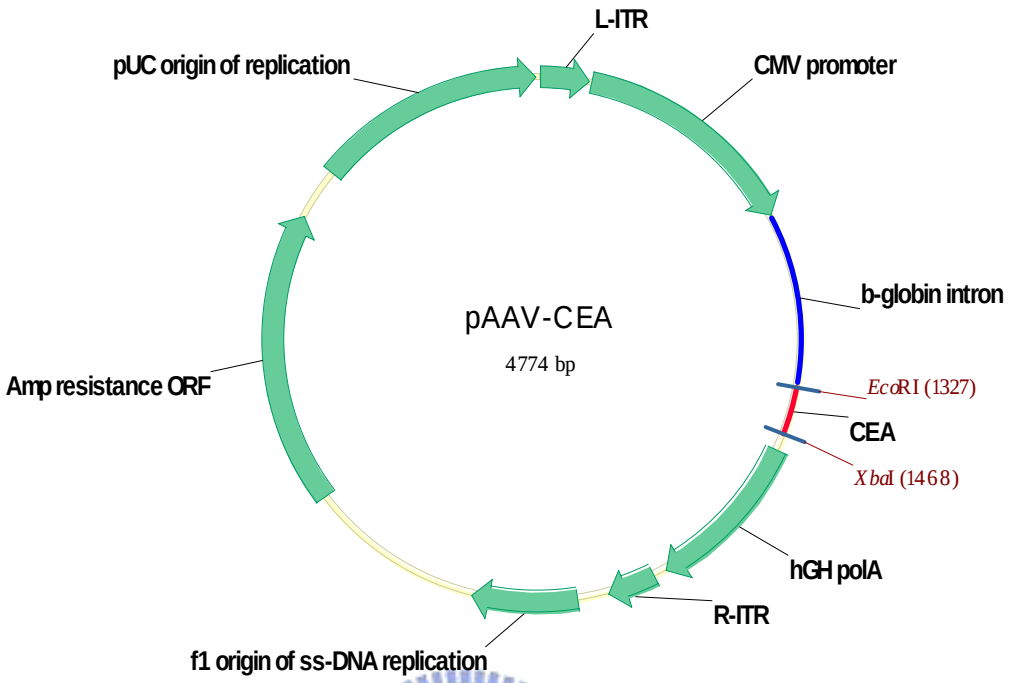
Xu, F., Y. J. Ye, et al. (2004). "In vitro antitumor immune response induced by fusion of dendritic cells and colon cancer cells." World J Gastroenterol **10**(8): 1162-6.

Yang, J.-M. (2004). "Development and evaluation of a generic evolutionary method for protein-ligand docking." Journal of Computational Chemistry **25**: 843-857.

Yang, J.-M. S., T.-W. (2005). "A pharmacohphore-based evoluntionary approach for screening selective estrogen receptor modulators." Proteins: Structure, Function, and Bioinformatics **59**: 205-220.

Zhang, Y., J. Li, et al. (2004). "Analysis of serum cytokines in patients with severe acute respiratory syndrome." Infect Immun **72**(8): 4410-5.

Appendix



Feature	Nucleotide Position
Left AAV-2 inverted terminal repeat (ITR)	1-141
CMV promoter	150-812
β-globin intron	820-1312
CEA	1332-1466
Human growth hormone (hGH) polyA signal	1519-1997
Right AAV-2 inverted terminal repeat (ITR)	2037-2177
fl origin of ss-DNA replication	2269-2575
Ampicillin resistance (*bla*) ORF	3094-3951
pUC origin of replication	4102-4769

```
   1  CCTGCAGGCA GCTGCGCGCT CGCTCGCTCA CTGAGGCCGC CCGGGCAAAG CCCGGGCGTC GGGCGACCTT
      GGACGTCCGT CGACGCGCGA GCGAGCGAGT GACTCCGGCG GGCCCGTTTC GGGCCCGCAG CCCGCTGGAA
  71  TGGTCGCCCG GCCTCAGTGA GCGAGCGAGC GCGCAGAGAG GGAGTGGCCA ACTCCATCAC TAGGGGTTCC
      ACCAGCGGGC CGGAGTCACT CGCTCGCTCG CGCGTCTCTC CCTCACCGGT TGAGGTAGTG ATCCCCAAGG
 141  TGCGGCCGCA CGCGTGGAGC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
      ACGCCGGCGT GCGCACCTCG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT
 211  TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCCGCCCATT
      ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA
 281  GACGTCAATA ATGACGTATG TTCCCATAGT AACGTCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG
      CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCAGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC
 351  TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC CCTATTGACG
      ATAAATGCCA TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCATGCGGG GGATAACTGC
 421  TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA
      AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT ACCCTGAAAG GATGAACCGT
 491  GTACATCTAC GTATTAGTCA TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA
      CATGTAGATG CATAATCAGT AGCGATAATG GTACCACTAC GCCAAAACCG TCATGTAGTT ACCCGCACCT
 561  TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGCACCA
      ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC AAAACGTGGT
 631  AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC AAATGGGCGG TAGGCGTGTA
      TTTAGTTGCC CTGAAAGGTT TTACAGCATT GTTGAGGCGG GGTAACTGCG TTTACCCGCC ATCCGCACAT
 701  CGGTGGGAGG TCTATATAAG CAGAGCTCGT TTAGTGAACC GTCAGATCGC CTGGAGACGC CATCCACGCT
      GCCACCCTCC AGATATATTC GTCTCGAGCA AATCACTTGG CAGTCTAGCG GACCTCTGCG GTAGGTGCGA
 771  GTTTTGACCT CCATAGAAGA CACCGGGACC GATCCAGCCT CCGCGGATTC GAATCCCGGC CGGGAACGGT
      CAAAACTGGA GGTATCTTCT GTGGCCCTGG CTAGGTCGGA GGCGCCTAAG CTTAGGGCCG GCCCTTGCCA
 841  GCATTGGAAC GCGGATTCCC CGTGCCAAGA GTGACGTAAG TACCGCCTAT AGAGTCTATA GGCCCACAAA
      CGTAACCTTG CGCCTAAGGG GCACGGTTCT CACTGCATTC ATGGCGGATA TCTCAGATAT CCGGGTGTTT
 911  AAATGCTTTC TTCTTTTAAT ATACTTTTTT GTTTATCTTA TTTCTAATAC TTTCCCTAAT CTCTTTCTTT
      TTTACGAAAG AAGAAAATTA TATGAAAAAA CAAATAGAAT AAAGATTATG AAAGGGATTA GAGAAAGAAA
 981  CAGGGCAATA ATGATACAAT GTATCATGCC TCTTTGCACC ATTCTAAAGA ATAACAGTGA TAATTTCTGG
      GTCCCGTTAT TACTATGTTA CATAGTACGG AGAAACGTGG TAAGATTTCT TATTGTCACT ATTAAAGACC
1051  GTTAAGGCAA TAGCAATATT TCTGCATATA AATATTTCTG CATATAAATT GTAACTGATG TAAGAGGTTT
      CAATTCCGTT ATCGTTATAA AGACGTATAT TTATAAAGAC GTATATTTAA CATTGACTAC ATTCTCCAAA
1121  CATATTGCTA ATAGCAGCTA CAATCCAGCT ACCATTCTGC TTTTATTTTA TGGTTGGGAT AAGGCTGGAT
      GTATAACGAT TATCGTCGAT GTTAGGTCGA TGGTAAGACG AAAATAAAAT ACCAACCCTA TTCCGACCTA
1191  TATTCTGAGT CCAAGCTAGG CCCTTTTGCT AATCATGTTC ATACCTCTTA TCTTCCTCCC ACAGCTCCTG
      ATAAGACTCA GGTTCGATCC GGGAAAACGA TTAGTACAAG TATGGAGAAT AGAAGGAGGG TGTCGAGGAC
1261  GGCAACGTGC TGGTCTGTGT GCTGGCCCAT CACTTTGGCA AAGAATTGGG ATTCGAACAT CGATTGAATT
      CCGTTGCACG ACCAGACACA CGACCGGGTA GTGAAACCGT TTCTTAACCC TAAGCTTGTA GCTAACTTAA
1331  CATGGAGTCT CCCTCGGCCC CTCCCCACAG ATGGTGCATC CCCTGGCAGA GGCTCCTGCT CACAGGTGAA
```

GTACCTCAGA GGGAGCCGGG GAGGGGTGTC TACCACGTAG GGGACCGTCT CCGAGGACGA GTGTCCACTT

1401 GGGAGGACAA CCTGGGAGAG GGTGGGAGGA GGGAGCTGGG GTCTCCTGGG TAGGACAGGG CTGTGATCTA

CCCTCCTGTT GGACCCTCTC CCACCCTCCT CCCTCGACCC CAGAGGACCC ATCCTGTCCC GACACTAGAT

1471 GAGTCGACCT GCAGAAGCTT GCCTCGAGCA GCGCTGCTCG AGAGATCTAC GGGTGGCATC CCTGTGACCC

CTCAGCTGGA CGTCTTCGAA CGGAGCTCGT CGCGACGAGC TCTCTAGATG CCCACCGTAG GGACACTGGG

1541 CTCCCCAGTG CCTCTCCTGG CCCTGGAAGT TGCCACTCCA GTGCCCACCA GCCTTGTCCT AATAAAATTA

GAGGGGTCAC GGAGAGGACC GGGACCTTCA ACGGTGAGGT CACGGGTGGT CGGAACAGGA TTATTTTAAT

1611 AGTTGCATCA TTTTGTCTGA CTAGGTGTCC TTCTATAATA TTATGGGGTG GAGGGGGGTG GTATGGAGCA

TCAACGTAGT AAAACAGACT GATCCACAGG AAGATATTAT AATACCCCAC CTCCCCCCAC CATACCTCGT

1681 AGGGGCAAGT TGGGAAGACA ACCTGTAGGG CCTGCGGGGT CTATTGGGAA CCAAGCTGGA GTGCAGTGGC

TCCCCGTTCA ACCCTTCTGT TGGACATCCC GGACGCCCCA GATAACCCTT GGTTCGACCT CACGTCACCG

1751 ACAATCTTGG CTCACTGCAA TCTCCGCCTC CTGGGTTCAA GCGATTCTCC TGCCTCAGCC TCCCGAGTTG

TGTTAGAACC GAGTGACGTT AGAGGCGGAG GACCCAAGTT CGCTAAGAGG ACGGAGTCGG AGGGCTCAAC

1821 TTGGGATTCC AGGCATGCAT GACCAGGCTC AGCTAATTTT TGTTTTTTTG GTAGAGACGG GGTTTCACCA

AACCCTAAGG TCCGTACGTA CTGGTCCGAG TCGATTAAAA ACAAAAAAAC CATCTCTGCC CCAAAGTGGT

1891 TATTGGCCAG GCTGGTCTCC AACTCCTAAT CTCAGGTGAT CTACCCACCT TGGCCTCCCA AATTGCTGGG

ATAACCGGTC CGACCAGAGG TTGAGGATTA GAGTCCACTA GATGGGTGGA ACCGGAGGGT TTAACGACCC

1961 ATTACAGGCG TGAACCACTG CTCCCTTCCC TGTCCTTCTG ATTTTGTAGG TAACCACGTG CGGACCGAGC

TAATGTCCGC ACTTGGTGAC GAGGGAAGGG ACAGGAAGAC TAAAACATCC ATTGGTGCAC GCCTGGCTCG

2031 GGCCGCAGGA ACCCCTAGTG ATGGAGTTGG CCACTCCCTC TCTGCGCGCT CGCTCGCTCA CTGAGGCCGG

CCGGCGTCCT TGGGGATCAC TACCTCAACC GGTGAGGGAG AGACGCGCGA GCGAGCGAGT GACTCCGGCC

2101 GCGACCAAAG GTCGCCCGAC GCCCGGGCTT TGCCCGGGCG GCCTCAGTGA GCGAGCGAGC GCGCAGCTGC

CGCTGGTTTC CAGCGGGCTG CGGGCCCGAA ACGGGCCCGC CGGAGTCACT CGCTCGCTCG CGCGTCGACG

2171 CTGCAGGGGC GCCTGATGCG GTATTTTCTC CTTACGCATC TGTGCGGTAT TTCACACCGC ATACGTCAAA

GACGTCCCCG CGGACTACGC CATAAAAGAG GAATGCGTAG ACACGCCATA AAGTGTGGCG TATGCAGTTT

2241 GCAACCATAG TACGCGCCCT GTAGCGGCGC ATTAAGCGCG GCGGGTGTGG TGGTTACGCG CAGCGTGACC

CGTTGGTATC ATGCGCGGGA CATCGCCGCG TAATTCGCGC CGCCCACACC ACCAATGCGC GTCGCACTGG

2311 GCTACACTTG CCAGCGCCCT AGCGCCCGCT CCTTTCGCTT TCTTCCCTTC CTTTCTCGCC ACGTTCGCCG

CGATGTGAAC GGTCGCGGGA TCGCGGGCGA GGAAAGCGAA AGAAGGGAAG GAAAGAGCGG TGCAAGCGGC

2381 GCTTTCCCCG TCAAGCTCTA AATCGGGGGC TCCCTTTAGG GTTCCGATTT AGTGCTTTAC GGCACCTCGA

CGAAAGGGGC AGTTCGAGAT TTAGCCCCCG AGGGAAATCC CAAGGCTAAA TCACGAAATG CCGTGGAGCT

2451 CCCCAAAAAA CTTGATTTGG GTGATGGTTC ACGTAGTGGG CCATCGCCCT GATAGACGGT TTTTCGCCCT

GGGGTTTTTT GAACTAAACC CACTACCAAG TGCATCACCC GGTAGCGGGA CTATCTGCCA AAAAGCGGGA

2521 TTGACGTTGG AGTCCACGTT CTTTAATAGT GGACTCTTGT TCCAAACTGG AACAACACTC AACCCTATCT

AACTGCAACC TCAGGTGCAA GAAATTATCA CCTGAGAACA AGGTTTGACC TTGTTGTGAG TTGGGATAGA

2591 CGGGCTATTC TTTTGATTTA TAAGGGATTT TGCCGATTTC GGCCTATTGG TTAAAAAATG AGCTGATTTA

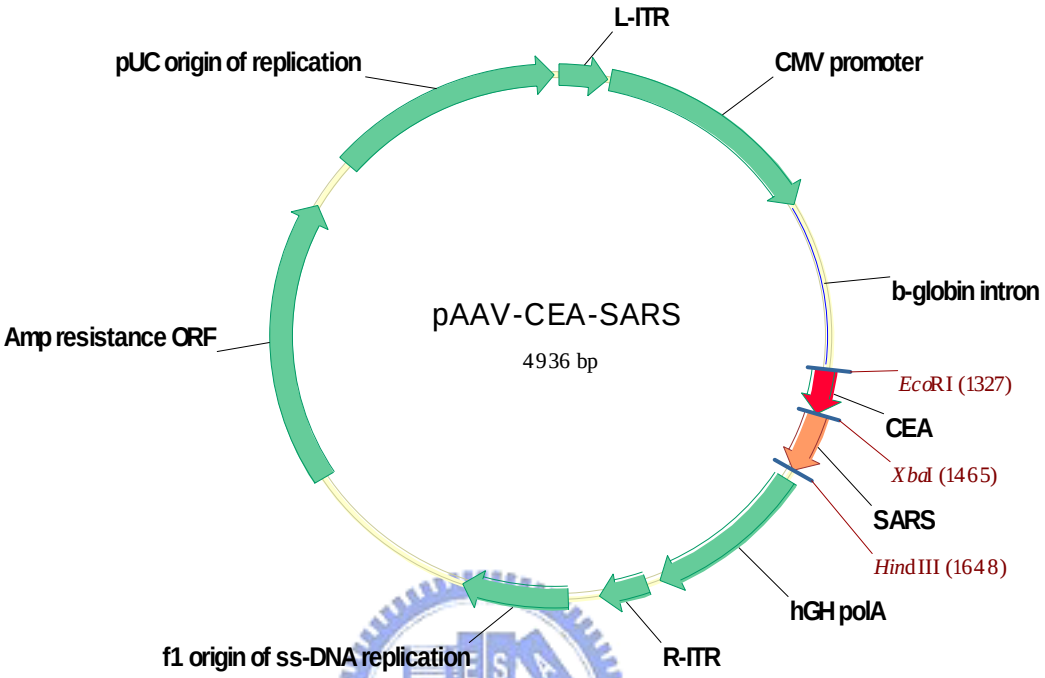
GCCCGATAAG AAAACTAAAT ATTCCCTAAA ACGGCTAAAG CCGGATAACC AATTTTTTAC TCGACTAAAT

2661 ACAAAAATTT AACGCGAATT TTAACAAAAT ATTAACGTTT ACAATTTTAT GGTGCACTCT CAGTACAATC

TGTTTTTAAA TTGCGCTTAA AATTGTTTTA TAATTGCAAA TGTTAAAATA CCACGTGAGA GTCATGTTAG

```
2731  TGCTCTGATG CCGCATAGTT AAGCCAGCCC CGACACCCGC CAACACCCGC TGACGCGCCC TGACGGGCTT
      ACGAGACTAC GGCGTATCAA TTCGGTCGGG GCTGTGGGCG GTTGTGGGCG ACTGCGCGGG ACTGCCCGAA
2801  GTCTGCTCCC GGCATCCGCT TACAGACAAG CTGTGACCGT CTCCGGGAGC TGCATGTGTC AGAGGTTTTC
      CAGACGAGGG CCGTAGGCGA ATGTCTGTTC GACACTGGCA GAGGCCCTCG ACGTACACAG TCTCCAAAAG
2871  ACCGTCATCA CCGAAACGCG CGAGACGAAA GGGCCTCGTG ATACGCCTAT TTTTATAGGT TAATGTCATG
      TGGCAGTAGT GGCTTTGCGC GCTCTGCTTT CCCGGAGCAC TATGCGGATA AAAATATCCA ATTACAGTAC
2941  ATAATAATGG TTTCTTAGAC GTCAGGTGGC ACTTTTCGGG GAAATGTGCG CGGAACCCCT ATTTGTTTAT
      TATTATTACC AAAGAATCTG CAGTCCACCG TGAAAAGCCC CTTTACACGC GCCTTGGGGA TAAACAAATA
3011  TTTTCTAAAT ACATTCAAAT ATGTATCCGC TCATGAGACA ATAACCCTGA TAAATGCTTC AATAATATTG
      AAAAGATTTA TGTAAGTTTA TACATAGGCG AGTACTCTGT TATTGGGACT ATTTACGAAG TTATTATAAC
3081  AAAAAGGAAG AGTATGAGTA TTCAACATTT CCGTGTCGCC CTTATTCCCT TTTTTGCGGC ATTTTGCCTT
      TTTTTCCTTC TCATACTCAT AAGTTGTAAA GGCACAGCGG GAATAAGGGA AAAAACGCCG TAAAACGGAA
3151  CCTGTTTTTG CTCACCCAGA AACGCTGGTG AAAGTAAAAG ATGCTGAAGA TCAGTTGGGT GCACGAGTGG
      GGACAAAAAC GAGTGGGTCT TTGCGACCAC TTTCATTTTC TACGACTTCT AGTCAACCCA CGTGCTCACC
3221  GTTACATCGA ACTGGATCTC AACAGCGGTA AGATCCTTGA GAGTTTTCGC CCCGAAGAAC GTTTTCCAAT
      CAATGTAGCT TGACCTAGAG TTGTCGCCAT TCTAGGAACT CTCAAAAGCG GGGCTTCTTG CAAAAGGTTA
3291  GATGAGCACT TTTAAAGTTC TGCTATGTGG CGCGGTATTA TCCCGTATTG ACGCCGGGCA AGAGCAACTC
      CTACTCGTGA AAATTTCAAG ACGATACACC GCGCCATAAT AGGGCATAAC TGCGGCCCGT TCTCGTTGAG
3361  GGTCGCCGCA TACACTATTC TCAGAATGAC TTGGTTGAGT ACTCACCAGT CACAGAAAAG CATCTTACGG
      CCAGCGGCGT ATGTGATAAG AGTCTTACTG AACCAACTCA TGAGTGGTCA GTGTCTTTTC GTAGAATGCC
3431  ATGGCATGAC AGTAAGAGAA TTATGCAGTG CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT
      TACCGTACTG TCATTCTCTT AATACGTCAC GACGGTATTG GTACTCACTA TTGTGACGCC GGTTGAATGA
3501  TCTGACAACG ATCGGAGGAC CGAAGGAGCT AACCGCTTTT TTGCACAACA TGGGGGATCA TGTAACTCGC
      AGACTGTTGC TAGCCTCCTG GCTTCCTCGA TTGGCGAAAA AACGTGTTGT ACCCCCTAGT ACATTGAGCG
3571  CTTGATCGTT GGGAACCGGA GCTGAATGAA GCCATACCAA ACGACGAGCG TGACACCACG ATGCCTGTAG
      GAACTAGCAA CCCTTGGCCT CGACTTACTT CGGTATGGTT TGCTGCTCGC ACTGTGGTGC TACGGACATC
3641  CAATGGCAAC AACGTTGCGC AAACTATTAA CTGGCGAACT ACTTACTCTA GCTTCCCGGC AACAATTAAT
      GTTACCGTTG TTGCAACGCG TTTGATAATT GACCGCTTGA TGAATGAGAT CGAAGGGCCG TTGTTAATTA
3711  AGACTGGATG GAGGCGGATA AAGTTGCAGG ACCACTTCTG CGCTCGGCCC TTCCGGCTGG CTGGTTTATT
      TCTGACCTAC CTCCGCCTAT TTCAACGTCC TGGTGAAGAC GCGAGCCGGG AAGGCCGACC GACCAAATAA
3781  GCTGATAAAT CTGGAGCCGG TGAGCGTGGG TCTCGCGGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC
      CGACTATTTA GACCTCGGCC ACTCGCACCC AGAGCGCCAT AGTAACGTCG TGACCCCGGT CTACCATTCG
3851  CCTCCCGTAT CGTAGTTATC TACACGACGG GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC
      GGAGGGCATA GCATCAATAG ATGTGCTGCC CCTCAGTCCG TTGATACCTA CTTGCTTTAT CTGTCTAGCG
3921  TGAGATAGGT GCCTCACTGA TTAAGCATTG GTAACTGTCA GACCAAGTTT ACTCATATAT ACTTTAGATT
      ACTCTATCCA CGGAGTGACT AATTCGTAAC CATTGACAGT CTGGTTCAAA TGAGTATATA TGAAATCTAA
3991  GATTTAAAAC TTCATTTTTA ATTTAAAAGG ATCTAGGTGA AGATCCTTTT TGATAATCTC ATGACCAAAA
      CTAAATTTTG AAGTAAAAAT TAAATTTTCC TAGATCCACT TCTAGGAAAA ACTATTAGAG TACTGGTTTT
4061  TCCCTTAACG TGAGTTTTCG TTCCACTGAG CGTCAGACCC CGTAGAAAAG ATCAAAGGAT CTTCTTGAGA
```

```
      AGGGAATTGC ACTCAAAAGC AAGGTGACTC GCAGTCTGGG GCATCTTTTC TAGTTTCCTA GAAGAACTCT
4131  TCCTTTTTTT CTGCGCGTAA TCTGCTGCTT GCAAACAAAA AAACCACCGC TACCAGCGGT GGTTTGTTTG
      AGGAAAAAAA GACGCGCATT AGACGACGAA CGTTTGTTTT TTTGGTGGCG ATGGTCGCCA CCAAACAAAC
4201  CCGGATCAAG AGCTACCAAC TCTTTTTCCG AAGGTAACTG GCTTCAGCAG AGCGCAGATA CCAAATACTG
      GGCCTAGTTC TCGATGGTTG AGAAAAAGGC TTCCATTGAC CGAAGTCGTC TCGCGTCTAT GGTTTATGAC
4271  TCCTTCTAGT GTAGCCGTAG TTAGGCCACC ACTTCAAGAA CTCTGTAGCA CCGCCTACAT ACCTCGCTCT
      AGGAAGATCA CATCGGCATC AATCCGGTGG TGAAGTTCTT GAGACATCGT GGCGGATGTA TGGAGCGAGA
4341  GCTAATCCTG TTACCAGTGG CTGCTGCCAG TGGCGATAAG TCGTGTCTTA CCGGGTTGGA CTCAAGACGA
      CGATTAGGAC AATGGTCACC GACGACGGTC ACCGCTATTC AGCACAGAAT GGCCCAACCT GAGTTCTGCT
4411  TAGTTACCGG ATAAGGCGCA GCGGTCGGGC TGAACGGGGG GTTCGTGCAC ACAGCCCAGC TTGGAGCGAA
      ATCAATGGCC TATTCCGCGT CGCCAGCCCG ACTTGCCCCC CAAGCACGTG TGTCGGGTCG AACCTCGCTT
4481  CGACCTACAC CGAACTGAGA TACCTACAGC GTGAGCTATG AGAAAGCGCC ACGCTTCCCG AAGGGAGAAA
      GCTGGATGTG GCTTGACTCT ATGGATGTCG CACTCGATAC TCTTTCGCGG TGCGAAGGGC TTCCCTCTTT
4551  GGCGGACAGG TATCCGGTAA GCGGCAGGGT CGGAACAGGA GAGCGCACGA GGGAGCTTCC AGGGGGAAAC
      CCGCCTGTCC ATAGGCCATT CGCCGTCCCA GCCTTGTCCT CTCGCGTGCT CCCTCGAAGG TCCCCCTTTG
4621  GCCTGGTATC TTTATAGTCC TGTCGGGTTT CGCCACCTCT GACTTGAGCG TCGATTTTTG TGATGCTCGT
      CGGACCATAG AAATATCAGG ACAGCCCAAA GCGGTGGAGA CTGAACTCGC AGCTAAAAAC ACTACGAGCA
4691  CAGGGGGGCG GAGCCTATGG AAAAACGCCA GCAACGCGGC CTTTTTACGG TTCCTGGCCT TTTGCTGGCC
      GTCCCCCCGC CTCGGATACC TTTTTGCGGT CGTTGCGCCG GAAAAATGCC AAGGACCGGA AAACGACCGG
4761  TTTTGCTCAC ATGT
      AAAACGAGTG TACA
```



Feature	Nucleotide Position
Left AAV-2 inverted terminal repeat (ITR)	1-141
CMV promoter	150-812
β-globin intron	820-1312
CEA	1332-1463
SARS	1469-1646
Human growth hormone (hGH) polyA signal	1681-2159
Right AAV-2 inverted terminal repeat (ITR)	2199-2339
fl origin of ss-DNA replication	2431-2737
Ampicillin resistance (*bla*) ORF	3256-4113
pUC origin of replication	4264-4931

```
   1  CCTGCAGGCA GCTGCGCGCT CGCTCGCTCA CTGAGGCCGC CCGGGCAAAG CCCGGGCGTC GGGCGACCTT
      GGACGTCCGT CGACGCGCGA GCGAGCGAGT GACTCCGGCG GGCCCGTTTC GGGCCCGCAG CCCGCTGGAA
  71  TGGTCGCCCG GCCTCAGTGA GCGAGCGAGC GCGCAGAGAG GGAGTGGCCA ACTCCATCAC TAGGGGTTCC
      ACCAGCGGGC CGGAGTCACT CGCTCGCTCG CGCGTCTCTC CCTCACCGGT TGAGGTAGTG ATCCCCAAGG
 141  TGCGGCCGCA CGCGTGGAGC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
      ACGCCGGCGT GCGCACCTCG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT
 211  TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCCGCCCATT
      ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA
 281  GACGTCAATA ATGACGTATG TTCCCATAGT AACGTCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG
      CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCAGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC
 351  TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC CCTATTGACG
      ATAAATGCCA TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCATGCGGG GGATAACTGC
 421  TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA
      AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT ACCCTGAAAG GATGAACCGT
 491  GTACATCTAC GTATTAGTCA TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA
      CATGTAGATG CATAATCAGT AGCGATAATG GTACCACTAC GCCAAAACCG TCATGTAGTT ACCCGCACCT
 561  TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGCACCA
      ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC AAAACGTGGT
 631  AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC AAATGGGCGG TAGGCGTGTA
      TTTAGTTGCC CTGAAAGGTT TTACAGCATT GTTGAGGCGG GGTAACTGCG TTTACCCGCC ATCCGCACAT
 701  CGGTGGGAGG TCTATATAAG CAGAGCTCGT TTAGTGAACC GTCAGATCGC CTGGAGACGC CATCCACGCT
      GCCACCCTCC AGATATATTC GTCTCGAGCA AATCACTTGG CAGTCTAGCG GACCTCTGCG GTAGGTGCGA
 771  GTTTTGACCT CCATAGAAGA CACCGGGACC GATCCAGCCT CCGCGGATTC GAATCCCGGC CGGGAACGGT
      CAAAACTGGA GGTATCTTCT GTGGCCCTGG CTAGGTCGGA GGCGCCTAAG CTTAGGGCCG GCCCTTGCCA
 841  GCATTGGAAC GCGGATTCCC CGTGCCAAGA GTGACGTAAG TACCGCCTAT AGAGTCTATA GGCCCACAAA
      CGTAACCTTG CGCCTAAGGG GCACGGTTCT CACTGCATTC ATGGCGGATA TCTCAGATAT CCGGGTGTTT
 911  AAATGCTTTC TTCTTTTAAT ATACTTTTTT GTTTATCTTA TTTCTAATAC TTTCCCTAAT CTCTTTCTTT
      TTTACGAAAG AAGAAAATTA TATGAAAAAA CAAATAGAAT AAAGATTATG AAAGGGATTA GAGAAAGAAA
 981  CAGGGCAATA ATGATACAAT GTATCATGCC TCTTTGCACC ATTCTAAAGA ATAACAGTGA TAATTTCTGG
      GTCCCGTTAT TACTATGTTA CATAGTACGG AGAAACGTGG TAAGATTTCT TATTGTCACT ATTAAAGACC
1051  GTTAAGGCAA TAGCAATATT TCTGCATATA AATATTTCTG CATATAAATT GTAACTGATG TAAGAGGTTT
      CAATTCCGTT ATCGTTATAA AGACGTATAT TTATAAAGAC GTATATTTAA CATTGACTAC ATTCTCCAAA
1121  CATATTGCTA ATAGCAGCTA CAATCCAGCT ACCATTCTGC TTTTATTTTA TGGTTGGGAT AAGGCTGGAT
      GTATAACGAT TATCGTCGAT GTTAGGTCGA TGGTAAGACG AAAATAAAAT ACCAACCCTA TTCCGACCTA
1191  TATTCTGAGT CCAAGCTAGG CCCTTTTGCT AATCATGTTC ATACCTCTTA TCTTCCTCCC ACAGCTCCTG
      ATAAGACTCA GGTTCGATCC GGGAAAACGA TTAGTACAAG TATGGAGAAT AGAAGGAGGG TGTCGAGGAC
1261  GGCAACGTGC TGGTCTGTGT GCTGGCCCAT CACTTTGGCA AAGAATTGGG ATTCGAACAT CGATTGAATT
      CCGTTGCACG ACCAGACACA CGACCGGGTA GTGAAACCGT TTCTTAACCC TAAGCTTGTA GCTAACTTAA
1331  CATGGAGTCT CCCTCGGCCC CTCCCCACAG ATGGTGCATC CCCTGGCAGA GGCTCCTGCT CACAGGTGAA
```

GTACCTCAGA GGGAGCCGGG GAGGGGTGTC TACCACGTAG GGGACCGTCT CCGAGGACGA GTGTCCACTT

1401 GGGAGGACAA CCTGGGAGAG GGTGGGAGGA GGGAGCTGGG GTCTCCTGGG TAGGACAGGG CTGTCTAGAA

CCCTCCTGTT GGACCCTCTC CCACCCTCCT CCCTCGACCC CAGAGGACCC ATCCTGTCCC GACAGATCTT

1471 AAGTCGAGGC GGAGGTACAA ATTGACAGGT TAATTACAGG CAGACTTCAA AGCCTTCAAA CCTATGTAAC

TTCAGCTCCG CCTCCATGTT TAACTGTCCA ATTAATGTCC GTCTGAAGTT TCGGAAGTTT GGATACATTG

1541 ACAACAACTA ATCAGGATTA AATGGCCTTG GTATGTTTGG CTCGGCTTCA TTGCTGGACT AATTGCCATC

TGTTGTTGAT TAGTCCTAAT TTACCGGAAC CATACAAACC GAGCCGAAGT AACGACCTGA TTAACGGTAG

1611 GTCATGGTTA CAATCTTGCT TTGTTGCATG ACTTAAAAGC TTGCCTCGAG CAGCGCTGCT CGAGAGATCT

CAGTACCAAT GTTAGAACGA AACAACGTAC TGAATTTTCG AACGGAGCTC GTCGCGACGA GCTCTCTAGA

1681 ACGGGTGGCA TCCCTGTGAC CCCTCCCCAG TGCCTCTCCT GGCCCTGGAA GTTGCCACTC CAGTGCCCAC

TGCCCACCGT AGGGACACTG GGGAGGGGTC ACGGAGAGGA CCGGGACCTT CAACGGTGAG GTCACGGGTG

1751 CAGCCTTGTC CTAATAAAAT TAAGTTGCAT CATTTTGTCT GACTAGGTGT CCTTCTATAA TATTATGGGG

GTCGGAACAG GATTATTTTA ATTCAACGTA GTAAAACAGA CTGATCCACA GGAAGATATT ATAATACCCC

1821 TGGAGGGGGG TGGTATGGAG CAAGGGGCAA GTTGGGAAGA CAACCTGTAG GGCCTGCGGG GTCTATTGGG

ACCTCCCCCC ACCATACCTC GTTCCCCGTT CAACCCTTCT GTTGGACATC CCGGACGCCC CAGATAACCC

1891 AACCAAGCTG GAGTGCAGTG GCACAATCTT GGCTCACTGC AATCTCCGCC TCCTGGGTTC AAGCGATTCT

TTGGTTCGAC CTCACGTCAC CGTGTTAGAA CCGAGTGACG TTAGAGGCGG AGGACCCAAG TTCGCTAAGA

1961 CCTGCCTCAG CCTCCCGAGT TGTTGGGATT CCAGGCATGC ATGACCAGGC TCAGCTAATT TTTGTTTTTT

GGACGGAGTC GGAGGGCTCA ACAACCCTAA GGTCCGTACG TACTGGTCCG AGTCGATTAA AAACAAAAAA

2031 TGGTAGAGAC GGGGTTTCAC CATATTGGCC AGGCTGGTCT CCAACTCCTA ATCTCAGGTG ATCTACCCAC

ACCATCTCTG CCCCAAAGTG GTATAACCGG TCCGACCAGA GGTTGAGGAT TAGAGTCCAC TAGATGGGTG

2101 CTTGGCCTCC CAAATTGCTG GGATTACAGG CGTGAACCAC TGCTCCCTTC CCTGTCCTTC TGATTTTGTA

GAACCGGAGG GTTTAACGAC CCTAATGTCC GCACTTGGTG ACGAGGGAAG GGACAGGAAG ACTAAAACAT

2171 GGTAACCACG TGCGGACCGA GCGGCCGCAG GAACCCCTAG TGATGGAGTT GGCCACTCCC TCTCTGCGCG

CCATTGGTGC ACGCCTGGCT CGCCGGCGTC CTTGGGGATC ACTACCTCAA CCGGTGAGGG AGAGACGCGC

2241 CTCGCTCGCT CACTGAGGCC GGGCGACCAA AGGTCGCCCG ACGCCCGGGC TTTGCCCGGG CGGCCTCAGT

GAGCGAGCGA GTGACTCCGG CCCGCTGGTT TCCAGCGGGC TGCGGGCCCG AAACGGGCCC GCCGGAGTCA

2311 GAGCGAGCGA GCGCGCAGCT GCCTGCAGGG GCGCCTGATG CGGTATTTTC TCCTTACGCA TCTGTGCGGT

CTCGCTCGCT CGCGCGTCGA CGGACGTCCC CGCGGACTAC GCCATAAAAG AGGAATGCGT AGACACGCCA

2381 ATTTCACACC GCATACGTCA AAGCAACCAT AGTACGCGCC CTGTAGCGGC GCATTAAGCG CGGCGGGTGT

TAAAGTGTGG CGTATGCAGT TTCGTTGGTA TCATGCGCGG GACATCGCCG CGTAATTCGC GCCGCCCACA

2451 GGTGGTTACG CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCTTTCGC TTTCTTCCCT

CCACCAATGC GCGTCGCACT GGCGATGTGA ACGGTCGCGG GATCGCGGGC GAGGAAAGCG AAAGAAGGGA

2521 TCCTTTCTCG CCACGTTCGC CGGCTTTCCC CGTCAAGCTC TAAATCGGGG GCTCCCTTTA GGGTTCCGAT

AGGAAAGAGC GGTGCAAGCG GCCGAAAGGG GCAGTTCGAG ATTTAGCCCC CGAGGGAAAT CCCAAGGCTA

2591 TTAGTGCTTT ACGGCACCTC GACCCCAAAA AACTTGATTT GGGTGATGGT TCACGTAGTG GGCCATCGCC

AATCACGAAA TGCCGTGGAG CTGGGGTTTT TTGAACTAAA CCCACTACCA AGTGCATCAC CCGGTAGCGG

2661 CTGATAGACG GTTTTTCGCC CTTTGACGTT GGAGTCCACG TTCTTTAATA GTGGACTCTT GTTCCAAACT

GACTATCTGC CAAAAAGCGG GAAACTGCAA CCTCAGGTGC AAGAAATTAT CACCTGAGAA CAAGGTTTGA

2731 GGAACAACAC TCAACCCTAT CTCGGGCTAT TCTTTTGATT TATAAGGGAT TTTGCCGATT TCGGCCTATT

CCTTGTTGTG AGTTGGGATA GAGCCCGATA AGAAAACTAA ATATTCCCTA AAACGGCTAA AGCCGGATAA

2801 GGTTAAAAAA TGAGCTGATT TAACAAAAAT TTAACGCGAA TTTTAACAAA ATATTAACGT TTACAATTTT

CCAATTTTTT ACTCGACTAA ATTGTTTTTA AATTGCGCTT AAAATTGTTT TATAATTGCA AATGTTAAAA

2871 ATGGTGCACT CTCAGTACAA TCTGCTCTGA TGCCGCATAG TTAAGCCAGC CCCGACACCC GCCAACACCC

TACCACGTGA GAGTCATGTT AGACGAGACT ACGGCGTATC AATTCGGTCG GGGCTGTGGG CGGTTGTGGG

2941 GCTGACGCGC CCTGACGGGC TTGTCTGCTC CCGGCATCCG CTTACAGACA AGCTGTGACC GTCTCCGGGA

CGACTGCGCG GGACTGCCCG AACAGACGAG GGCCGTAGGC GAATGTCTGT TCGACACTGG CAGAGGCCCT

3011 GCTGCATGTG TCAGAGGTTT TCACCGTCAT CACCGAAACG CGCGAGACGA AAGGGCCTCG TGATACGCCT

CGACGTACAC AGTCTCCAAA AGTGGCAGTA GTGGCTTTGC GCGCTCTGCT TTCCCGGAGC ACTATGCGGA

3081 ATTTTTATAG GTTAATGTCA TGATAATAAT GGTTTCTTAG ACGTCAGGTG GCACTTTTCG GGGAAATGTG

TAAAAATATC CAATTACAGT ACTATTATTA CCAAAGAATC TGCAGTCCAC CGTGAAAAGC CCCTTTACAC

3151 CGCGGAACCC CTATTTGTTT ATTTTTCTAA ATACATTCAA ATATGTATCC GCTCATGAGA CAATAACCCT

GCGCCTTGGG GATAAACAAA TAAAAAGATT TATGTAAGTT TATACATAGG CGAGTACTCT GTTATTGGGA

3221 GATAAATGCT TCAATAATAT TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCG CCCTTATTCC

CTATTTACGA AGTTATTATA ACTTTTTCCT TCTCATACTC ATAAGTTGTA AAGGCACAGC GGGAATAAGG

3291 CTTTTTTGCG GCATTTTGCC TTCCTGTTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA AGATGCTGAA

GAAAAAACGC CGTAAAACGG AAGGACAAAA ACGAGTGGGT CTTTGCGACC ACTTTCATTT TCTACGACTT

3361 GATCAGTTGG GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG TAAGATCCTT GAGAGTTTTC

CTAGTCAACC CACGTGCTCA CCCAATGTAG CTTGACCTAG AGTTGTCGCC ATTCTAGGAA CTCTCAAAAG

3431 GCCCCGAAGA ACGTTTTCCA ATGATGAGCA CTTTTAAAGT TCTGCTATGT GGCGCGGTAT TATCCCGTAT

CGGGGCTTCT TGCAAAAGGT TACTACTCGT GAAAATTTCA AGACGATACA CCGCGCCATA ATAGGGCATA

3501 TGACGCCGGG CAAGAGCAAC TCGGTCGCCG CATACACTAT TCTCAGAATG ACTTGGTTGA GTACTCACCA

ACTGCGGCCC GTTCTCGTTG AGCCAGCGGC GTATGTGATA AGAGTCTTAC TGAACCAACT CATGAGTGGT

3571 GTCACAGAAA AGCATCTTAC GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGAGTG

CAGTGTCTTT TCGTAGAATG CCTACCGTAC TGTCATTCTC TTAATACGTC ACGACGGTAT TGGTACTCAC

3641 ATAACACTGC GGCCAACTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTTGCACAA

TATTGTGACG CCGGTTGAAT GAAGACTGTT GCTAGCCTCC TGGCTTCCTC GATTGGCGAA AAAACGTGTT

3711 CATGGGGGAT CATGTAACTC GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC AAACGACGAG

GTACCCCCTA GTACATTGAG CGGAACTAGC AACCCTTGGC CTCGACTTAC TTCGGTATGG TTTGCTGCTC

3781 CGTGACACCA CGATGCCTGT AGCAATGGCA ACAACGTTGC GCAAACTATT AACTGGCGAA CTACTTACTC

GCACTGTGGT GCTACGGACA TCGTTACCGT TGTTGCAACG CGTTTGATAA TTGACCGCTT GATGAATGAG

3851 TAGCTTCCCG GCAACAATTA ATAGACTGGA TGGAGGCGGA TAAAGTTGCA GGACCACTTC TGCGCTCGGC

ATCGAAGGGC CGTTGTTAAT TATCTGACCT ACCTCCGCCT ATTTCAACGT CCTGGTGAAG ACGCGAGCCG

3921 CCTTCCGGCT GGCTGGTTTA TTGCTGATAA ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG TATCATTGCA

GGAAGGCCGA CCGACCAAAT AACGACTATT TAGACCTCGG CCACTCGCAC CCAGAGCGCC ATAGTAACGT

3991 GCACTGGGGC CAGATGGTAA GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG GCAACTATGG

CGTGACCCCG GTCTACCATT CGGGAGGGCA TAGCATCAAT AGATGTGCTG CCCCTCAGTC CGTTGATACC

4061 ATGAACGAAA TAGACAGATC GCTGAGATAG GTGCCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT

TACTTGCTTT ATCTGTCTAG CGACTCTATC CACGGAGTGA CTAATTCGTA ACCATTGACA GTCTGGTTCA

4131 TTACTCATAT ATACTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAAA GGATCTAGGT GAAGATCCTT

AATGAGTATA TATGAAATCT AACTAAATTT TGAAGTAAAA ATTAAATTTT CCTAGATCCA CTTCTAGGAA

4201 TTTGATAATC TCATGACCAA AATCCCTTAA CGTGAGTTTT CGTTCCACTG AGCGTCAGAC CCCGTAGAAA

AAACTATTAG AGTACTGGTT TTAGGGAATT GCACTCAAAA GCAAGGTGAC TCGCAGTCTG GGGCATCTTT

4271 AGATCAAAGG ATCTTCTTGA GATCCTTTTT TTCTGCGCGT AATCTGCTGC TTGCAAACAA AAAAACCACC

TCTAGTTTCC TAGAAGAACT CTAGGAAAAA AAGACGCGCA TTAGACGACG AACGTTTGTT TTTTTGGTGG

4341 GCTACCAGCG GTGGTTTGTT TGCCGGATCA AGAGCTACCA ACTCTTTTTC CGAAGGTAAC TGGCTTCAGC

CGATGGTCGC CACCAAACAA ACGGCCTAGT TCTCGATGGT TGAGAAAAAG GCTTCCATTG ACCGAAGTCG

4411 AGAGCGCAGA TACCAAATAC TGTCCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG

TCTCGCGTCT ATGGTTTATG ACAGGAAGAT CACATCGGCA TCAATCCGGT GGTGAAGTTC TTGAGACATC

4481 CACCGCCTAC ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT

GTGGCGGATG TATGGAGCGA GACGATTAGG ACAATGGTCA CCGACGACGG TCACCGCTAT TCAGCACAGA

4551 TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG GCTGAACGGG GGGTTCGTGC

ATGGCCCAAC CTGAGTTCTG CTATCAATGG CCTATTCCGC GTCGCCAGCC CGACTTGCCC CCCAAGCACG

4621 ACACAGCCCA GCTTGGAGCG AACGACCTAC ACCGAACTGA GATACCTACA GCGTGAGCTA TGAGAAAGCG

TGTGTCGGGT CGAACCTCGC TTGCTGGATG TGGCTTGACT CTATGGATGT CGCACTCGAT ACTCTTTCGC

4691 CCACGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG GAGAGCGCAC

GGTGCGAAGG GCTTCCCTCT TTCCGCCTGT CCATAGGCCA TTCGCCGTCC CAGCCTTGTC CTCTCGCGTG

4761 GAGGGAGCTT CCAGGGGGAA ACGCCTGGTA TCTTTATAGT CCTGTCGGGT TTCGCCACCT CTGACTTGAG

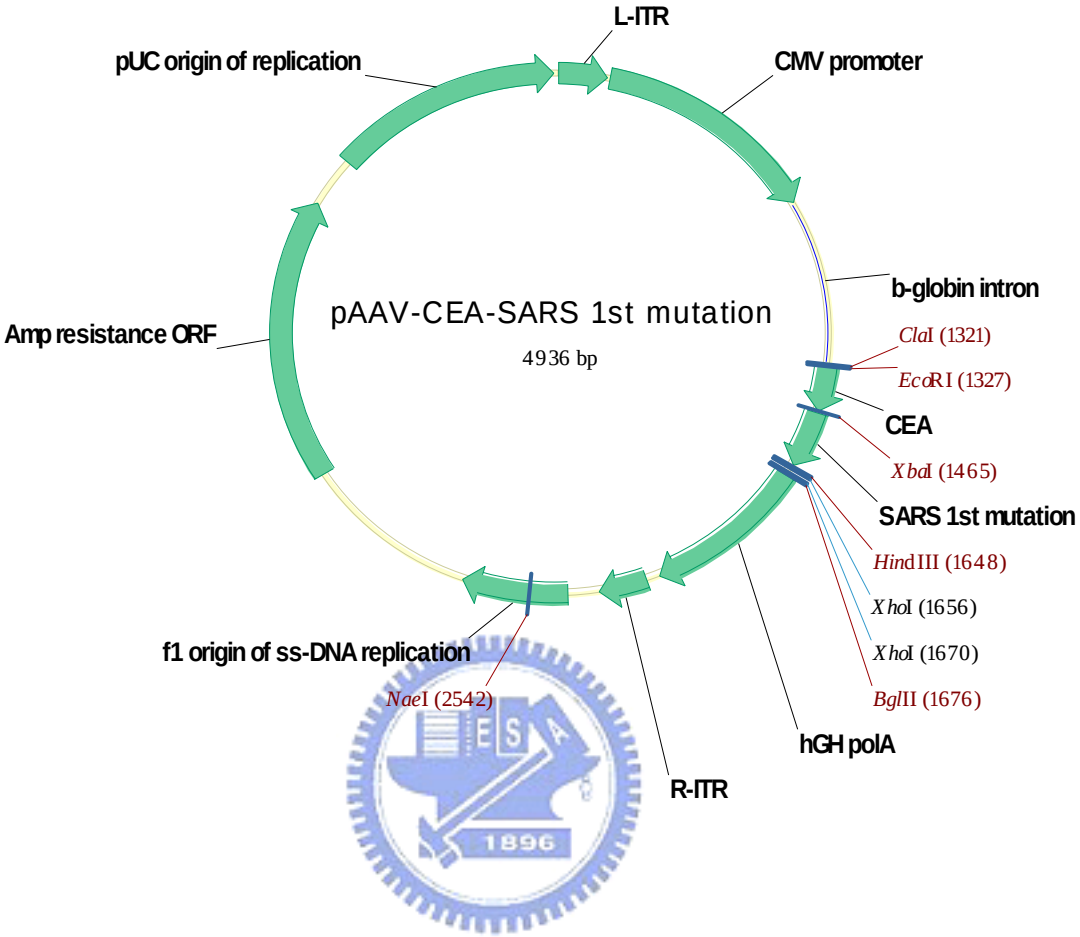
CTCCCTCGAA GGTCCCCCTT TGCGGACCAT AGAAATATCA GGACAGCCCA AAGCGGTGGA GACTGAACTC

4831 CGTCGATTTT TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG GCCTTTTTAC

GCAGCTAAAA ACACTACGAG CAGTCCCCCC GCCTCGGATA CCTTTTTGCG GTCGTTGCGC CGGAAAAATG

4901 GGTTCCTGGC CTTTTGCTGG CCTTTTGCTC ACATGT

CCAAGGACCG GAAAACGACC GGAAAACGAG TGTACA



Feature	Nucleotide Position
Left AAV-2 inverted terminal repeat (ITR)	1-141
CMV promoter	150-812
β-globin intron	820-1312
CEA	1332-1463
SARS	1469-1646
Human growth hormone (hGH) polyA signal	1681-2159
Right AAV-2 inverted terminal repeat (ITR)	2199-2339
fl origin of ss-DNA replication	2431-2737
Ampicillin resistance (*bla*) ORF	3256-4113
pUC origin of replication	4264-4931

```
   1  CCTGCAGGCA GCTGCGCGCT CGCTCGCTCA CTGAGGCCGC CCGGGCAAAG CCCGGGCGTC GGGCGACCTT
      GGACGTCCGT CGACGCGCGA GCGAGCGAGT GACTCCGGCG GGCCCGTTTC GGGCCCGCAG CCCGCTGGAA
  71  TGGTCGCCCG GCCTCAGTGA GCGAGCGAGC GCGCAGAGAG GGAGTGGCCA ACTCCATCAC TAGGGGTTCC
      ACCAGCGGGC CGGAGTCACT CGCTCGCTCG CGCGTCTCTC CCTCACCGGT TGAGGTAGTG ATCCCCAAGG
 141  TGCGGCCGCA CGCGTGGAGC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
      ACGCCGGCGT GCGCACCTCG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT
 211  TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCCGCCCATT
      ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA
 281  GACGTCAATA ATGACGTATG TTCCCATAGT AACGTCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG
      CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCAGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC
 351  TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC CCTATTGACG
      ATAAATGCCA TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCATGCGGG GGATAACTGC
 421  TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA
      AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT ACCCTGAAAG GATGAACCGT
 491  GTACATCTAC GTATTAGTCA TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA
      CATGTAGATG CATAATCAGT AGCGATAATG GTACCACTAC GCCAAAACCG TCATGTAGTT ACCCGCACCT
 561  TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGCACCA
      ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC AAAACGTGGT
 631  AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC AAATGGGCGG TAGGCGTGTA
      TTTAGTTGCC CTGAAAGGTT TTACAGCATT GTTGAGGCGG GGTAACTGCG TTTACCCGCC ATCCGCACAT
 701  CGGTGGGAGG TCTATATAAG CAGAGCTCGT TTAGTGAACC GTCAGATCGC CTGGAGACGC CATCCACGCT
      GCCACCCTCC AGATATATTC GTCTCGAGCA AATCACTTGG CAGTCTAGCG GACCTCTGCG GTAGGTGCGA
 771  GTTTTGACCT CCATAGAAGA CACCGGGACC GATCCAGCCT CCGCGGATTC GAATCCCGGC CGGGAACGGT
      CAAAACTGGA GGTATCTTCT GTGGCCCTGG CTAGGTCGGA GGCGCCTAAG CTTAGGGCCG GCCCTTGCCA
 841  GCATTGGAAC GCGGATTCCC CGTGCCAAGA GTGACGTAAG TACCGCCTAT AGAGTCTATA GGCCCACAAA
      CGTAACCTTG CGCCTAAGGG GCACGGTTCT CACTGCATTC ATGGCGGATA TCTCAGATAT CCGGGTGTTT
 911  AAATGCTTTC TTCTTTTAAT ATACTTTTTT GTTTATCTTA TTTCTAATAC TTTCCCTAAT CTCTTTCTTT
      TTTACGAAAG AAGAAAATTA TATGAAAAAA CAAATAGAAT AAAGATTATG AAAGGGATTA GAGAAAGAAA
 981  CAGGGCAATA ATGATACAAT GTATCATGCC TCTTTGCACC ATTCTAAAGA ATAACAGTGA TAATTTCTGG
      GTCCCGTTAT TACTATGTTA CATAGTACGG AGAAACGTGG TAAGATTTCT TATTGTCACT ATTAAAGACC
1051  GTTAAGGCAA TAGCAATATT TCTGCATATA AATATTTCTG CATATAAATT GTAACTGATG TAAGAGGTTT
      CAATTCCGTT ATCGTTATAA AGACGTATAT TTATAAAGAC GTATATTTAA CATTGACTAC ATTCTCCAAA
1121  CATATTGCTA ATAGCAGCTA CAATCCAGCT ACCATTCTGC TTTTATTTTA TGGTTGGGAT AAGGCTGGAT
      GTATAACGAT TATCGTCGAT GTTAGGTCGA TGGTAAGACG AAAATAAAAT ACCAACCCTA TTCCGACCTA
1191  TATTCTGAGT CCAAGCTAGG CCCTTTTGCT AATCATGTTC ATACCTCTTA TCTTCCTCCC ACAGCTCCTG
      ATAAGACTCA GGTTCGATCC GGGAAAACGA TTAGTACAAG TATGGAGAAT AGAAGGAGGG TGTCGAGGAC
1261  GGCAACGTGC TGGTCTGTGT GCTGGCCCAT CACTTTGGCA AAGAATTGGG ATTCGAACAT CGATTGAATT
      CCGTTGCACG ACCAGACACA CGACCGGGTA GTGAAACCGT TTCTTAACCC TAAGCTTGTA GCTAACTTAA
1331  CATGGAGTCT CCCTCGGCCC CTCCCCACAG ATGGTGCATC CCCTGGCAGA GGCTCCTGCT CACAGGTGAA
```

GTACCTCAGA GGGAGCCGGG GAGGGGTGTC TACCACGTAG GGGACCGTCT CCGAGGACGA GTGTCCACTT

1401 GGGAGGACAA CCTGGGAGAG GGTGGGAGGA GGGAGCTGGG GTCTCCTGGG TAGGACAGGG CTGTCTAGAA

CCCTCCTGTT GGACCCTCTC CCACCCTCCT CCCTCGACCC CAGAGGACCC ATCCTGTCCC GACAGATCTT

1471 AAGTCGAGGC GGAGGTACAA ATTGACAGGT TAATTACAGG CAGACTTCAA AGCCTTCAAA CCTATGTAAC

TTCAGCTCCG CCTCCATGTT TAACTGTCCA ATTAATGTCC GTCTGAAGTT TCGGAAGTTT GGATACATTG

1541 ACAACAACTA ATCAGGATTA AATGGCCTTG GTATGTTTGG CTCGGCTTCA TTATTGGACT AATTGCCATC

TGTTGTTGAT TAGTCCTAAT TTACCGGAAC CATACAAACC GAGCCGAAGT AATAACCTGA TTAACGGTAG

1611 GTCATGGTTA CAATCTTGCT TTGTTGCATG ACTTAAAAGC TTGCCTCGAG CAGCGCTGCT CGAGAGATCT

CAGTACCAAT GTTAGAACGA AACAACGTAC TGAATTTTCG AACGGAGCTC GTCGCGACGA GCTCTCTAGA

1681 ACGGGTGGCA TCCCTGTGAC CCCTCCCCAG TGCCTCTCCT GGCCCTGGAA GTTGCCACTC CAGTGCCCAC

TGCCCACCGT AGGGACACTG GGGAGGGGTC ACGGAGAGGA CCGGGACCTT CAACGGTGAG GTCACGGGTG

1751 CAGCCTTGTC CTAATAAAAT TAAGTTGCAT CATTTTGTCT GACTAGGTGT CCTTCTATAA TATTATGGGG

GTCGGAACAG GATTATTTTA ATTCAACGTA GTAAAACAGA CTGATCCACA GGAAGATATT ATAATACCCC

1821 TGGAGGGGGG TGGTATGGAG CAAGGGGCAA GTTGGGAAGA CAACCTGTAG GGCCTGCGGG GTCTATTGGG

ACCTCCCCCC ACCATACCTC GTTCCCCGTT CAACCCTTCT GTTGGACATC CCGGACGCCC CAGATAACCC

1891 AACCAAGCTG GAGTGCAGTG GCACAATCTT GGCTCACTGC AATCTCCGCC TCCTGGGTTC AAGCGATTCT

TTGGTTCGAC CTCACGTCAC CGTGTTAGAA CCGAGTGACG TTAGAGGCGG AGGACCCAAG TTCGCTAAGA

1961 CCTGCCTCAG CCTCCCGAGT TGTTGGGATT CCAGGCATGC ATGACCAGGC TCAGCTAATT TTTGTTTTTT

GGACGGAGTC GGAGGGCTCA ACAACCCTAA GGTCCGTACG TACTGGTCCG AGTCGATTAA AAACAAAAAA

2031 TGGTAGAGAC GGGGTTTCAC CATATTGGCC AGGCTGGTCT CCAACTCCTA ATCTCAGGTG ATCTACCCAC

ACCATCTCTG CCCCAAAGTG GTATAACCGG TCCGACCAGA GGTTGAGGAT TAGAGTCCAC TAGATGGGTG

2101 CTTGGCCTCC CAAATTGCTG GGATTACAGG CGTGAACCAC TGCTCCCTTC CCTGTCCTTC TGATTTTGTA

GAACCGGAGG GTTTAACGAC CCTAATGTCC GCACTTGGTG ACGAGGGAAG GGACAGGAAG ACTAAAACAT

2171 GGTAACCACG TGCGGACCGA GCGGCCGCAG GAACCCCTAG TGATGGAGTT GGCCACTCCC TCTCTGCGCG

CCATTGGTGC ACGCCTGGCT CGCCGGCGTC CTTGGGGATC ACTACCTCAA CCGGTGAGGG AGAGACGCGC

2241 CTCGCTCGCT CACTGAGGCC GGGCGACCAA AGGTCGCCCG ACGCCCGGGC TTTGCCCGGG CGGCCTCAGT

GAGCGAGCGA GTGACTCCGG CCCGCTGGTT TCCAGCGGGC TGCGGGCCCG AAACGGGCCC GCCGGAGTCA

2311 GAGCGAGCGA GCGCGCAGCT GCCTGCAGGG GCGCCTGATG CGGTATTTTC TCCTTACGCA TCTGTGCGGT

CTCGCTCGCT CGCGCGTCGA CGGACGTCCC CGCGGACTAC GCCATAAAAG AGGAATGCGT AGACACGCCA

2381 ATTTCACACC GCATACGTCA AAGCAACCAT AGTACGCGCC CTGTAGCGGC GCATTAAGCG CGGCGGGTGT

TAAAGTGTGG CGTATGCAGT TTCGTTGGTA TCATGCGCGG GACATCGCCG CGTAATTCGC GCCGCCCACA

2451 GGTGGTTACG CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCTTTCGC TTTCTTCCCT

CCACCAATGC GCGTCGCACT GGCGATGTGA ACGGTCGCGG GATCGCGGGC GAGGAAAGCG AAAGAAGGGA

2521 TCCTTTCTCG CCACGTTCGC CGGCTTTCCC CGTCAAGCTC TAAATCGGGG GCTCCCTTTA GGGTTCCGAT

AGGAAAGAGC GGTGCAAGCG GCCGAAAGGG GCAGTTCGAG ATTTAGCCCC CGAGGGAAAT CCCAAGGCTA

2591 TTAGTGCTTT ACGGCACCTC GACCCCAAAA AACTTGATTT GGGTGATGGT TCACGTAGTG GGCCATCGCC

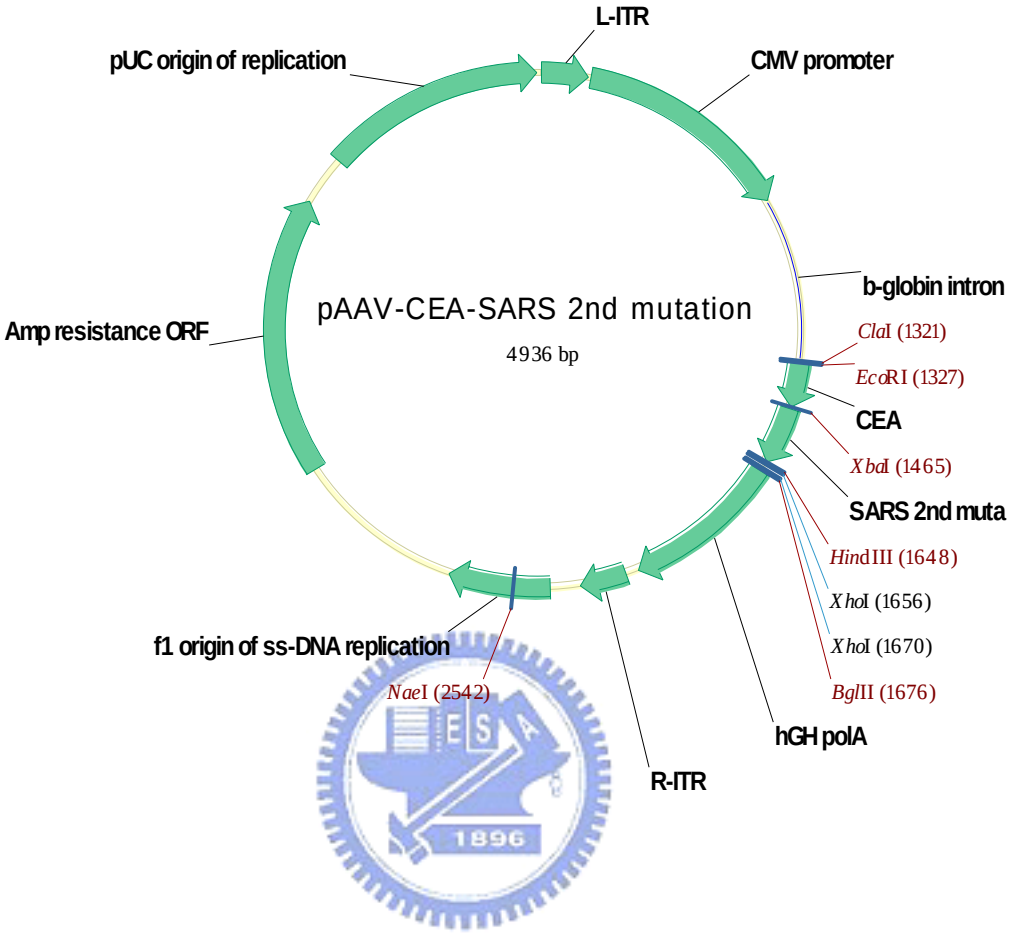
AATCACGAAA TGCCGTGGAG CTGGGGTTTT TTGAACTAAA CCCACTACCA AGTGCATCAC CCGGTAGCGG

2661 CTGATAGACG GTTTTTCGCC CTTTGACGTT GGAGTCCACG TTCTTTAATA GTGGACTCTT GTTCCAAACT

GACTATCTGC CAAAAAGCGG GAAACTGCAA CCTCAGGTGC AAGAAATTAT CACCTGAGAA CAAGGTTTGA

```
2731  GGAACAACAC TCAACCCTAT CTCGGGCTAT TCTTTTGATT TATAAGGGAT TTTGCCGATT TCGGCCTATT
      CCTTGTTGTG AGTTGGGATA GAGCCCGATA AGAAAACTAA ATATTCCCTA AAACGGCTAA AGCCGGATAA
2801  GGTTAAAAAA TGAGCTGATT TAACAAAAAT TTAACGCGAA TTTTAACAAA ATATTAACGT TTACAATTTT
      CCAATTTTTT ACTCGACTAA ATTGTTTTTA AATTGCGCTT AAAATTGTTT TATAATTGCA AATGTTAAAA
2871  ATGGTGCACT CTCAGTACAA TCTGCTCTGA TGCCGCATAG TTAAGCCAGC CCCGACACCC GCCAACACCC
      TACCACGTGA GAGTCATGTT AGACGAGACT ACGGCGTATC AATTCGGTCG GGGCTGTGGG CGGTTGTGGG
2941  GCTGACGCGC CCTGACGGGC TTGTCTGCTC CCGGCATCCG CTTACAGACA AGCTGTGACC GTCTCCGGGA
      CGACTGCGCG GGACTGCCCG AACAGACGAG GGCCGTAGGC GAATGTCTGT TCGACACTGG CAGAGGCCCT
3011  GCTGCATGTG TCAGAGGTTT TCACCGTCAT CACCGAAACG CGCGAGACGA AAGGGCCTCG TGATACGCCT
      CGACGTACAC AGTCTCCAAA AGTGGCAGTA GTGGCTTTGC GCGCTCTGCT TTCCCGGAGC ACTATGCGGA
3081  ATTTTTATAG GTTAATGTCA TGATAATAAT GGTTTCTTAG ACGTCAGGTG GCACTTTTCG GGGAAATGTG
      TAAAAATATC CAATTACAGT ACTATTATTA CCAAAGAATC TGCAGTCCAC CGTGAAAAGC CCCTTTACAC
3151  CGCGGAACCC CTATTTGTTT ATTTTTCTAA ATACATTCAA ATATGTATCC GCTCATGAGA CAATAACCCT
      GCGCCTTGGG GATAAACAAA TAAAAAGATT TATGTAAGTT TATACATAGG CGAGTACTCT GTTATTGGGA
3221  GATAAATGCT TCAATAATAT TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCG CCCTTATTCC
      CTATTTACGA AGTTATTATA ACTTTTTCCT TCTCATACTC ATAAGTTGTA AAGGCACAGC GGGAATAAGG
3291  CTTTTTTGCG GCATTTTGCC TTCCTGTTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA AGATGCTGAA
      GAAAAAACGC CGTAAAACGG AAGGACAAAA ACGAGTGGGT CTTTGCGACC ACTTTCATTT TCTACGACTT
3361  GATCAGTTGG GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG TAAGATCCTT GAGAGTTTTC
      CTAGTCAACC CACGTGCTCA CCCAATGTAG CTTGACCTAG AGTTGTCGCC ATTCTAGGAA CTCTCAAAAG
3431  GCCCCGAAGA ACGTTTTCCA ATGATGAGCA CTTTTAAAGT TCTGCTATGT GGCGCGGTAT TATCCCGTAT
      CGGGGCTTCT TGCAAAAGGT TACTACTCGT GAAAATTTCA AGACGATACA CCGCGCCATA ATAGGGCATA
3501  TGACGCCGGG CAAGAGCAAC TCGGTCGCCG CATACACTAT TCTCAGAATG ACTTGGTTGA GTACTCACCA
      ACTGCGGCCC GTTCTCGTTG AGCCAGCGGC GTATGTGATA AGAGTCTTAC TGAACCAACT CATGAGTGGT
3571  GTCACAGAAA AGCATCTTAC GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGAGTG
      CAGTGTCTTT TCGTAGAATG CCTACCGTAC TGTCATTCTC TTAATACGTC ACGACGGTAT TGGTACTCAC
3641  ATAACACTGC GGCCAACTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTTGCACAA
      TATTGTGACG CCGGTTGAAT GAAGACTGTT GCTAGCCTCC TGGCTTCCTC GATTGGCGAA AAAACGTGTT
3711  CATGGGGGAT CATGTAACTC GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC AAACGACGAG
      GTACCCCCTA GTACATTGAG CGGAACTAGC AACCCTTGGC CTCGACTTAC TTCGGTATGG TTTGCTGCTC
3781  CGTGACACCA CGATGCCTGT AGCAATGGCA ACAACGTTGC GCAAACTATT AACTGGCGAA CTACTTACTC
      GCACTGTGGT GCTACGGACA TCGTTACCGT TGTTGCAACG CGTTTGATAA TTGACCGCTT GATGAATGAG
3851  TAGCTTCCCG GCAACAATTA ATAGACTGGA TGGAGGCGGA TAAAGTTGCA GGACCACTTC TGCGCTCGGC
      ATCGAAGGGC CGTTGTTAAT TATCTGACCT ACCTCCGCCT ATTTCAACGT CCTGGTGAAG ACGCGAGCCG
3921  CCTTCCGGCT GGCTGGTTTA TTGCTGATAA ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG TATCATTGCA
      GGAAGGCCGA CCGACCAAAT AACGACTATT TAGACCTCGG CCACTCGCAC CCAGAGCGCC ATAGTAACGT
3991  GCACTGGGGC CAGATGGTAA GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG GCAACTATGG
      CGTGACCCCG GTCTACCATT CGGGAGGGCA TAGCATCAAT AGATGTGCTG CCCCTCAGTC CGTTGATACC
4061  ATGAACGAAA TAGACAGATC GCTGAGATAG GTGCCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT
```

```
       TACTTGCTTT ATCTGTCTAG CGACTCTATC CACGGAGTGA CTAATTCGTA ACCATTGACA GTCTGGTTCA
4131   TTACTCATAT ATACTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAAA GGATCTAGGT GAAGATCCTT
       AATGAGTATA TATGAAATCT AACTAAATTT TGAAGTAAAA ATTAAATTTT CCTAGATCCA CTTCTAGGAA
4201   TTTGATAATC TCATGACCAA AATCCCTTAA CGTGAGTTTT CGTTCCACTG AGCGTCAGAC CCCGTAGAAA
       AAACTATTAG AGTACTGGTT TTAGGGAATT GCACTCAAAA GCAAGGTGAC TCGCAGTCTG GGGCATCTTT
4271   AGATCAAAGG ATCTTCTTGA GATCCTTTTT TTCTGCGCGT AATCTGCTGC TTGCAAACAA AAAAACCACC
       TCTAGTTTCC TAGAAGAACT CTAGGAAAAA AAGACGCGCA TTAGACGACG AACGTTTGTT TTTTTGGTGG
4341   GCTACCAGCG GTGGTTTGTT TGCCGGATCA AGAGCTACCA ACTCTTTTTC CGAAGGTAAC TGGCTTCAGC
       CGATGGTCGC CACCAAACAA ACGGCCTAGT TCTCGATGGT TGAGAAAAAG GCTTCCATTG ACCGAAGTCG
4411   AGAGCGCAGA TACCAAATAC TGTCCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG
       TCTCGCGTCT ATGGTTTATG ACAGGAAGAT CACATCGGCA TCAATCCGGT GGTGAAGTTC TTGAGACATC
4481   CACCGCCTAC ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT
       GTGGCGGATG TATGGAGCGA GACGATTAGG ACAATGGTCA CCGACGACGG TCACCGCTAT TCAGCACAGA
4551   TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG GCTGAACGGG GGGTTCGTGC
       ATGGCCCAAC CTGAGTTCTG CTATCAATGG CCTATTCCGC GTCGCCAGCC CGACTTGCCC CCCAAGCACG
4621   ACACAGCCCA GCTTGGAGCG AACGACCTAC ACCGAACTGA GATACCTACA GCGTGAGCTA TGAGAAAGCG
       TGTGTCGGGT CGAACCTCGC TTGCTGGATG TGGCTTGACT CTATGGATGT CGCACTCGAT ACTCTTTCGC
4691   CCACGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG GAGAGCGCAC
       GGTGCGAAGG GCTTCCCTCT TTCCGCCTGT CCATAGGCCA TTCGCCGTCC CAGCCTTGTC CTCTCGCGTG
4761   GAGGGAGCTT CCAGGGGGAA ACGCCTGGTA TCTTTATAGT CCTGTCGGGT TTCGCCACCT CTGACTTGAG
       CTCCCTCGAA GGTCCCCCTT TGCGGACCAT AGAAATATCA GGACAGCCCA AAGCGGTGGA GACTGAACTC
4831   CGTCGATTTT TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG GCCTTTTTAC
       GCAGCTAAAA ACACTACGAG CAGTCCCCCC GCCTCGGATA CCTTTTTGCG GTCGTTGCGC CGGAAAAATG
4901   GGTTCCTGGC CTTTTGCTGG CCTTTTGCTC ACATGT
       CCAAGGACCG GAAAACGACC GGAAAACGAG TGTACA
```



Feature	Nucleotide Position
Left AAV-2 inverted terminal repeat (ITR)	1-141
CMV promoter	150-812
β-globin intron	820-1312
CEA	1332-1463
SARS	1469-1646
Human growth hormone (hGH) polyA signal	1681-2159
Right AAV-2 inverted terminal repeat (ITR)	2199-2339
fl origin of ss-DNA replication	2431-2737
Ampicillin resistance (*bla*) ORF	3256-4113
pUC origin of replication	4264-4931

```
   1  CCTGCAGGCA GCTGCGCGCT CGCTCGCTCA CTGAGGCCGC CCGGGCAAAG CCCGGGCGTC GGGCGACCTT
      GGACGTCCGT CGACGCGCGA GCGAGCGAGT GACTCCGGCG GGCCCGTTTC GGGCCCGCAG CCCGCTGGAA
  71  TGGTCGCCCG GCCTCAGTGA GCGAGCGAGC GCGCAGAGAG GGAGTGGCCA ACTCCATCAC TAGGGGTTCC
      ACCAGCGGGC CGGAGTCACT CGCTCGCTCG CGCGTCTCTC CCTCACCGGT TGAGGTAGTG ATCCCCAAGG
 141  TGCGGCCGCA CGCGTGGAGC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
      ACGCCGGCGT GCGCACCTCG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT
 211  TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCCGCCCATT
      ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA
 281  GACGTCAATA ATGACGTATG TTCCCATAGT AACGTCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG
      CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCAGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC
 351  TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC CCTATTGACG
      ATAAATGCCA TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCATGCGGG GGATAACTGC
 421  TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA
      AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT ACCCTGAAAG GATGAACCGT
 491  GTACATCTAC GTATTAGTCA TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA
      CATGTAGATG CATAATCAGT AGCGATAATG GTACCACTAC GCCAAAACCG TCATGTAGTT ACCCGCACCT
 561  TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGCACCA
      ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC AAAACGTGGT
 631  AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC AAATGGGCGG TAGGCGTGTA
      TTTAGTTGCC CTGAAAGGTT TTACAGCATT GTTGAGGCGG GGTAACTGCG TTTACCCGCC ATCCGCACAT
 701  CGGTGGGAGG TCTATATAAG CAGAGCTCGT TTAGTGAACC GTCAGATCGC CTGGAGACGC CATCCACGCT
      GCCACCCTCC AGATATATTC GTCTCGAGCA AATCACTTGG CAGTCTAGCG GACCTCTGCG GTAGGTGCGA
 771  GTTTTGACCT CCATAGAAGA CACCGGGACC GATCCAGCCT CCGCGGATTC GAATCCCGGC CGGGAACGGT
      CAAAACTGGA GGTATCTTCT GTGGCCCTGG CTAGGTCGGA GGCGCCTAAG CTTAGGGCCG GCCCTTGCCA
 841  GCATTGGAAC GCGGATTCCC CGTGCCAAGA GTGACGTAAG TACCGCCTAT AGAGTCTATA GGCCCACAAA
      CGTAACCTTG CGCCTAAGGG GCACGGTTCT CACTGCATTC ATGGCGGATA TCTCAGATAT CCGGGTGTTT
 911  AAATGCTTTC TTCTTTTAAT ATACTTTTTT GTTTATCTTA TTTCTAATAC TTTCCCTAAT CTCTTTCTTT
      TTTACGAAAG AAGAAAATTA TATGAAAAAA CAAATAGAAT AAAGATTATG AAAGGGATTA GAGAAAGAAA
 981  CAGGGCAATA ATGATACAAT GTATCATGCC TCTTTGCACC ATTCTAAAGA ATAACAGTGA TAATTTCTGG
      GTCCCGTTAT TACTATGTTA CATAGTACGG AGAAACGTGG TAAGATTTCT TATTGTCACT ATTAAAGACC
1051  GTTAAGGCAA TAGCAATATT TCTGCATATA AATATTTCTG CATATAAATT GTAACTGATG TAAGAGGTTT
      CAATTCCGTT ATCGTTATAA AGACGTATAT TTATAAAGAC GTATATTTAA CATTGACTAC ATTCTCCAAA
1121  CATATTGCTA ATAGCAGCTA CAATCCAGCT ACCATTCTGC TTTTATTTTA TGGTTGGGAT AAGGCTGGAT
      GTATAACGAT TATCGTCGAT GTTAGGTCGA TGGTAAGACG AAAATAAAAT ACCAACCCTA TTCCGACCTA
1191  TATTCTGAGT CCAAGCTAGG CCCTTTTGCT AATCATGTTC ATACCTCTTA TCTTCCTCCC ACAGCTCCTG
      ATAAGACTCA GGTTCGATCC GGGAAAACGA TTAGTACAAG TATGGAGAAT AGAAGGAGGG TGTCGAGGAC
1261  GGCAACGTGC TGGTCTGTGT GCTGGCCCAT CACTTTGGCA AAGAATTGGG ATTCGAACAT CGATTGAATT
      CCGTTGCACG ACCAGACACA CGACCGGGTA GTGAAACCGT TTCTTAACCC TAAGCTTGTA GCTAACTTAA
1331  CATGGAGTCT CCCTCGGCCC CTCCCCACAG ATGGTGCATC CCCTGGCAGA GGCTCCTGCT CACAGGTGAA
```

GTACCTCAGA GGGAGCCGGG GAGGGGTGTC TACCACGTAG GGGACCGTCT CCGAGGACGA GTGTCCACTT

1401 GGGAGGACAA CCTGGGAGAG GGTGGGAGGA GGGAGCTGGG GTCTCCTGGG TAGGACAGGG CTGTCTAGAA

CCCTCCTGTT GGACCCTCTC CCACCCTCCT CCCTCGACCC CAGAGGACCC ATCCTGTCCC GACAGATCTT

1471 AAGTCGAGGC GGAGGTACAA ATTGACAGGT TAATTACAGG CAGACTTCAA AGCCTTCAAA CCTATGTAAC

TTCAGCTCCG CCTCCATGTT TAACTGTCCA ATTAATGTCC GTCTGAAGTT TCGGAAGTTT GGATACATTG

1541 ACAACAACTA ATCAGGATTA AATGGCCTTG GTATGTTTGG CTCGGCACCA TTATTGGACT AATTGCCATC

TGTTGTTGAT TAGTCCTAAT TTACCGGAAC CATACAAACC GAGCCGTGGT AATAACCTGA TTAACGGTAG

1611 GTCATGGTTA CAATCTTGCT TTGTTGCATG ACTTAAAAGC TTGCCTCGAG CAGCGCTGCT CGAGAGATCT

CAGTACCAAT GTTAGAACGA AACAACGTAC TGAATTTTCG AACGGAGCTC GTCGCGACGA GCTCTCTAGA

1681 ACGGGTGGCA TCCCTGTGAC CCCTCCCCAG TGCCTCTCCT GGCCCTGGAA GTTGCCACTC CAGTGCCCAC

TGCCCACCGT AGGGACACTG GGGAGGGGTC ACGGAGAGGA CCGGGACCTT CAACGGTGAG GTCACGGGTG

1751 CAGCCTTGTC CTAATAAAAT TAAGTTGCAT CATTTTGTCT GACTAGGTGT CCTTCTATAA TATTATGGGG

GTCGGAACAG GATTATTTTA ATTCAACGTA GTAAAACAGA CTGATCCACA GGAAGATATT ATAATACCCC

1821 TGGAGGGGGG TGGTATGGAG CAAGGGGCAA GTTGGGAAGA CAACCTGTAG GGCCTGCGGG GTCTATTGGG

ACCTCCCCCC ACCATACCTC GTTCCCCGTT CAACCCTTCT GTTGGACATC CCGGACGCCC CAGATAACCC

1891 AACCAAGCTG GAGTGCAGTG GCACAATCTT GGCTCACTGC AATCTCCGCC TCCTGGGTTC AAGCGATTCT

TTGGTTCGAC CTCACGTCAC CGTGTTAGAA CCGAGTGACG TTAGAGGCGG AGGACCCAAG TTCGCTAAGA

1961 CCTGCCTCAG CCTCCCGAGT TGTTGGGATT CCAGGCATGC ATGACCAGGC TCAGCTAATT TTTGTTTTTT

GGACGGAGTC GGAGGGCTCA ACAACCCTAA GGTCCGTACG TACTGGTCCG AGTCGATTAA AAACAAAAAA

2031 TGGTAGAGAC GGGGTTTCAC CATATTGGCC AGGCTGGTCT CCAACTCCTA ATCTCAGGTG ATCTACCCAC

ACCATCTCTG CCCCAAAGTG GTATAACCGG TCCGACCAGA GGTTGAGGAT TAGAGTCCAC TAGATGGGTG

2101 CTTGGCCTCC CAAATTGCTG GGATTACAGG CGTGAACCAC TGCTCCCTTC CCTGTCCTTC TGATTTTGTA

GAACCGGAGG GTTTAACGAC CCTAATGTCC GCACTTGGTG ACGAGGGAAG GGACAGGAAG ACTAAAACAT

2171 GGTAACCACG TGCGGACCGA GCGGCCGCAG GAACCCCTAG TGATGGAGTT GGCCACTCCC TCTCTGCGCG

CCATTGGTGC ACGCCTGGCT CGCCGGCGTC CTTGGGGATC ACTACCTCAA CCGGTGAGGG AGAGACGCGC

2241 CTCGCTCGCT CACTGAGGCC GGGCGACCAA AGGTCGCCCG ACGCCCGGGC TTTGCCCGGG CGGCCTCAGT

GAGCGAGCGA GTGACTCCGG CCCGCTGGTT TCCAGCGGGC TGCGGGCCCG AAACGGGCCC GCCGGAGTCA

2311 GAGCGAGCGA GCGCGCAGCT GCCTGCAGGG GCGCCTGATG CGGTATTTTC TCCTTACGCA TCTGTGCGGT

CTCGCTCGCT CGCGCGTCGA CGGACGTCCC CGCGGACTAC GCCATAAAAG AGGAATGCGT AGACACGCCA

2381 ATTTCACACC GCATACGTCA AAGCAACCAT AGTACGCGCC CTGTAGCGGC GCATTAAGCG CGGCGGGTGT

TAAAGTGTGG CGTATGCAGT TTCGTTGGTA TCATGCGCGG GACATCGCCG CGTAATTCGC GCCGCCCACA

2451 GGTGGTTACG CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCTTTCGC TTTCTTCCCT

CCACCAATGC GCGTCGCACT GGCGATGTGA ACGGTCGCGG GATCGCGGGC GAGGAAAGCG AAAGAAGGGA

2521 TCCTTTCTCG CCACGTTCGC CGGCTTTCCC CGTCAAGCTC TAAATCGGGG GCTCCCTTTA GGGTTCCGAT

AGGAAAGAGC GGTGCAAGCG GCCGAAAGGG GCAGTTCGAG ATTTAGCCCC CGAGGGAAAT CCCAAGGCTA

2591 TTAGTGCTTT ACGGCACCTC GACCCCAAAA AACTTGATTT GGGTGATGGT TCACGTAGTG GGCCATCGCC

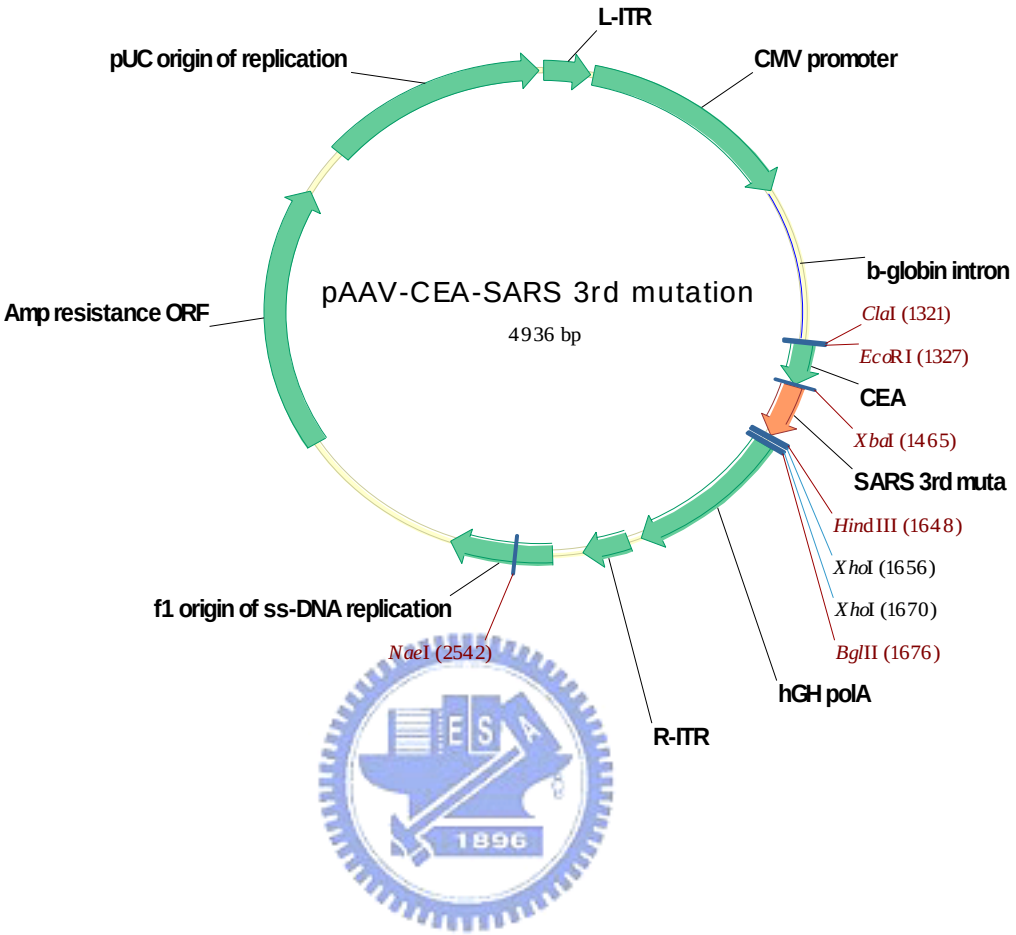
AATCACGAAA TGCCGTGGAG CTGGGGTTTT TTGAACTAAA CCCACTACCA AGTGCATCAC CCGGTAGCGG

2661 CTGATAGACG GTTTTTCGCC CTTTGACGTT GGAGTCCACG TTCTTTAATA GTGGACTCTT GTTCCAAACT

GACTATCTGC CAAAAAGCGG GAAACTGCAA CCTCAGGTGC AAGAAATTAT CACCTGAGAA CAAGGTTTGA

```
2731  GGAACAACAC TCAACCCTAT CTCGGGCTAT TCTTTTGATT TATAAGGGAT TTTGCCGATT TCGGCCTATT
      CCTTGTTGTG AGTTGGGATA GAGCCCGATA AGAAAACTAA ATATTCCCTA AAACGGCTAA AGCCGGATAA
2801  GGTTAAAAAA TGAGCTGATT TAACAAAAAT TTAACGCGAA TTTTAACAAA ATATTAACGT TTACAATTTT
      CCAATTTTTT ACTCGACTAA ATTGTTTTTA AATTGCGCTT AAAATTGTTT TATAATTGCA AATGTTAAAA
2871  ATGGTGCACT CTCAGTACAA TCTGCTCTGA TGCCGCATAG TTAAGCCAGC CCCGACACCC GCCAACACCC
      TACCACGTGA GAGTCATGTT AGACGAGACT ACGGCGTATC AATTCGGTCG GGGCTGTGGG CGGTTGTGGG
2941  GCTGACGCGC CCTGACGGGC TTGTCTGCTC CCGGCATCCG CTTACAGACA AGCTGTGACC GTCTCCGGGA
      CGACTGCGCG GGACTGCCCG AACAGACGAG GGCCGTAGGC GAATGTCTGT TCGACACTGG CAGAGGCCCT
3011  GCTGCATGTG TCAGAGGTTT TCACCGTCAT CACCGAAACG CGCGAGACGA AAGGGCCTCG TGATACGCCT
      CGACGTACAC AGTCTCCAAA AGTGGCAGTA GTGGCTTTGC GCGCTCTGCT TTCCCGGAGC ACTATGCGGA
3081  ATTTTTATAG GTTAATGTCA TGATAATAAT GGTTTCTTAG ACGTCAGGTG GCACTTTTCG GGGAAATGTG
      TAAAAATATC CAATTACAGT ACTATTATTA CCAAAGAATC TGCAGTCCAC CGTGAAAAGC CCCTTTACAC
3151  CGCGGAACCC CTATTTGTTT ATTTTTCTAA ATACATTCAA ATATGTATCC GCTCATGAGA CAATAACCCT
      GCGCCTTGGG GATAAACAAA TAAAAAGATT TATGTAAGTT TATACATAGG CGAGTACTCT GTTATTGGGA
3221  GATAAATGCT TCAATAATAT TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCG CCCTTATTCC
      CTATTTACGA AGTTATTATA ACTTTTTCCT TCTCATACTC ATAAGTTGTA AAGGCACAGC GGGAATAAGG
3291  CTTTTTTGCG GCATTTTGCC TTCCTGTTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA AGATGCTGAA
      GAAAAAACGC CGTAAAACGG AAGGACAAAA ACGAGTGGGT CTTTGCGACC ACTTTCATTT TCTACGACTT
3361  GATCAGTTGG GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG TAAGATCCTT GAGAGTTTTC
      CTAGTCAACC CACGTGCTCA CCCAATGTAG CTTGACCTAG AGTTGTCGCC ATTCTAGGAA CTCTCAAAAG
3431  GCCCCGAAGA ACGTTTTCCA ATGATGAGCA CTTTTAAAGT TCTGCTATGT GGCGCGGTAT TATCCCGTAT
      CGGGGCTTCT TGCAAAAGGT TACTACTCGT GAAAATTTCA AGACGATACA CCGCGCCATA ATAGGGCATA
3501  TGACGCCGGG CAAGAGCAAC TCGGTCGCCG CATACACTAT TCTCAGAATG ACTTGGTTGA GTACTCACCA
      ACTGCGGCCC GTTCTCGTTG AGCCAGCGGC GTATGTGATA AGAGTCTTAC TGAACCAACT CATGAGTGGT
3571  GTCACAGAAA AGCATCTTAC GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGAGTG
      CAGTGTCTTT TCGTAGAATG CCTACCGTAC TGTCATTCTC TTAATACGTC ACGACGGTAT TGGTACTCAC
3641  ATAACACTGC GGCCAACTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTTGCACAA
      TATTGTGACG CCGGTTGAAT GAAGACTGTT GCTAGCCTCC TGGCTTCCTC GATTGGCGAA AAAACGTGTT
3711  CATGGGGGAT CATGTAACTC GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC AAACGACGAG
      GTACCCCCTA GTACATTGAG CGGAACTAGC AACCCTTGGC CTCGACTTAC TTCGGTATGG TTTGCTGCTC
3781  CGTGACACCA CGATGCCTGT AGCAATGGCA ACAACGTTGC GCAAACTATT AACTGGCGAA CTACTTACTC
      GCACTGTGGT GCTACGGACA TCGTTACCGT TGTTGCAACG CGTTTGATAA TTGACCGCTT GATGAATGAG
3851  TAGCTTCCCG GCAACAATTA ATAGACTGGA TGGAGGCGGA TAAAGTTGCA GGACCACTTC TGCGCTCGGC
      ATCGAAGGGC CGTTGTTAAT TATCTGACCT ACCTCCGCCT ATTTCAACGT CCTGGTGAAG ACGCGAGCCG
3921  CCTTCCGGCT GGCTGGTTTA TTGCTGATAA ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG TATCATTGCA
      GGAAGGCCGA CCGACCAAAT AACGACTATT TAGACCTCGG CCACTCGCAC CCAGAGCGCC ATAGTAACGT
3991  GCACTGGGGC CAGATGGTAA GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG GCAACTATGG
      CGTGACCCCG GTCTACCATT CGGGAGGGCA TAGCATCAAT AGATGTGCTG CCCCTCAGTC CGTTGATACC
4061  ATGAACGAAA TAGACAGATC GCTGAGATAG GTGCCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT
```

```
      TACTTGCTTT ATCTGTCTAG CGACTCTATC CACGGAGTGA CTAATTCGTA ACCATTGACA GTCTGGTTCA
4131  TTACTCATAT ATACTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAAA GGATCTAGGT GAAGATCCTT
      AATGAGTATA TATGAAATCT AACTAAATTT TGAAGTAAAA ATTAAATTTT CCTAGATCCA CTTCTAGGAA
4201  TTTGATAATC TCATGACCAA AATCCCTTAA CGTGAGTTTT CGTTCCACTG AGCGTCAGAC CCCGTAGAAA
      AAACTATTAG AGTACTGGTT TTAGGGAATT GCACTCAAAA GCAAGGTGAC TCGCAGTCTG GGGCATCTTT
4271  AGATCAAAGG ATCTTCTTGA GATCCTTTTT TTCTGCGCGT AATCTGCTGC TTGCAAACAA AAAAACCACC
      TCTAGTTTCC TAGAAGAACT CTAGGAAAAA AAGACGCGCA TTAGACGACG AACGTTTGTT TTTTTGGTGG
4341  GCTACCAGCG GTGGTTTGTT TGCCGGATCA AGAGCTACCA ACTCTTTTTC CGAAGGTAAC TGGCTTCAGC
      CGATGGTCGC CACCAAACAA ACGGCCTAGT TCTCGATGGT TGAGAAAAAG GCTTCCATTG ACCGAAGTCG
4411  AGAGCGCAGA TACCAAATAC TGTCCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG
      TCTCGCGTCT ATGGTTTATG ACAGGAAGAT CACATCGGCA TCAATCCGGT GGTGAAGTTC TTGAGACATC
4481  CACCGCCTAC ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT
      GTGGCGGATG TATGGAGCGA GACGATTAGG ACAATGGTCA CCGACGACGG TCACCGCTAT TCAGCACAGA
4551  TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG GCTGAACGGG GGGTTCGTGC
      ATGGCCCAAC CTGAGTTCTG CTATCAATGG CCTATTCCGC GTCGCCAGCC CGACTTGCCC CCCAAGCACG
4621  ACACAGCCCA GCTTGGAGCG AACGACCTAC ACCGAACTGA GATACCTACA GCGTGAGCTA TGAGAAAGCG
      TGTGTCGGGT CGAACCTCGC TTGCTGGATG TGGCTTGACT CTATGGATGT CGCACTCGAT ACTCTTTCGC
4691  CCACGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG GAGAGCGCAC
      GGTGCGAAGG GCTTCCCTCT TTCCGCCTGT CCATAGGCCA TTCGCCGTCC CAGCCTTGTC CTCTCGCGTG
4761  GAGGGAGCTT CCAGGGGGAA ACGCCTGGTA TCTTTATAGT CCTGTCGGGT TTCGCCACCT CTGACTTGAG
      CTCCCTCGAA GGTCCCCCTT TGCGGACCAT AGAAATATCA GGACAGCCCA AAGCGGTGGA GACTGAACTC
4831  CGTCGATTTT TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG GCCTTTTTAC
      GCAGCTAAAA ACACTACGAG CAGTCCCCCC GCCTCGGATA CCTTTTTGCG GTCGTTGCGC CGGAAAAATG
4901  GGTTCCTGGC CTTTTGCTGG CCTTTTGCTC ACATGT
      CCAAGGACCG GAAAACGACC GGAAAACGAG TGTACA
```



Feature	Nucleotide Position
Left AAV-2 inverted terminal repeat (ITR)	1-141
CMV promoter	150-812
β-globin intron	820-1312
CEA	1332-1463
SARS	1469-1646
Human growth hormone (hGH) polyA signal	1681-2159
Right AAV-2 inverted terminal repeat (ITR)	2199-2339
fl origin of ss-DNA replication	2431-2737
Ampicillin resistance (*bla*) ORF	3256-4113
pUC origin of replication	4264-4931

```
   1  CCTGCAGGCA GCTGCGCGCT CGCTCGCTCA CTGAGGCCGC CCGGGCAAAG CCCGGGCGTC GGGCGACCTT
      GGACGTCCGT CGACGCGCGA GCGAGCGAGT GACTCCGGCG GGCCCGTTTC GGGCCCGCAG CCCGCTGGAA
  71  TGGTCGCCCG GCCTCAGTGA GCGAGCGAGC GCGCAGAGAG GGAGTGGCCA ACTCCATCAC TAGGGGTTCC
      ACCAGCGGGC CGGAGTCACT CGCTCGCTCG CGCGTCTCTC CCTCACCGGT TGAGGTAGTG ATCCCCAAGG
 141  TGCGGCCGCA CGCGTGGAGC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
      ACGCCGGCGT GCGCACCTCG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT
 211  TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCCGCCCATT
      ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA
 281  GACGTCAATA ATGACGTATG TTCCCATAGT AACGTCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG
      CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCAGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC
 351  TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC CCTATTGACG
      ATAAATGCCA TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCATGCGGG GGATAACTGC
 421  TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA
      AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT ACCCTGAAAG GATGAACCGT
 491  GTACATCTAC GTATTAGTCA TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA
      CATGTAGATG CATAATCAGT AGCGATAATG GTACCACTAC GCCAAAACCG TCATGTAGTT ACCCGCACCT
 561  TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGCACCA
      ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC AAAACGTGGT
 631  AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC AAATGGGCGG TAGGCGTGTA
      TTTAGTTGCC CTGAAAGGTT TTACAGCATT GTTGAGGCGG GGTAACTGCG TTTACCCGCC ATCCGCACAT
 701  CGGTGGGAGG TCTATATAAG CAGAGCTCGT TTAGTGAACC GTCAGATCGC CTGGAGACGC CATCCACGCT
      GCCACCCTCC AGATATATTC GTCTCGAGCA AATCACTTGG CAGTCTAGCG GACCTCTGCG GTAGGTGCGA
 771  GTTTTGACCT CCATAGAAGA CACCGGGACC GATCCAGCCT CCGCGGATTC GAATCCCGGC CGGGAACGGT
      CAAAACTGGA GGTATCTTCT GTGGCCCTGG CTAGGTCGGA GGCGCCTAAG CTTAGGGCCG GCCCTTGCCA
 841  GCATTGGAAC GCGGATTCCC CGTGCCAAGA GTGACGTAAG TACCGCCTAT AGAGTCTATA GGCCCACAAA
      CGTAACCTTG CGCCTAAGGG GCACGGTTCT CACTGCATTC ATGGCGGATA TCTCAGATAT CCGGGTGTTT
 911  AAATGCTTTC TTCTTTTAAT ATACTTTTTT GTTTATCTTA TTTCTAATAC TTTCCCTAAT CTCTTTCTTT
      TTTACGAAAG AAGAAAATTA TATGAAAAAA CAAATAGAAT AAAGATTATG AAAGGGATTA GAGAAAGAAA
 981  CAGGGCAATA ATGATACAAT GTATCATGCC TCTTTGCACC ATTCTAAAGA ATAACAGTGA TAATTTCTGG
      GTCCCGTTAT TACTATGTTA CATAGTACGG AGAAACGTGG TAAGATTTCT TATTGTCACT ATTAAAGACC
1051  GTTAAGGCAA TAGCAATATT TCTGCATATA AATATTTCTG CATATAAATT GTAACTGATG TAAGAGGTTT
      CAATTCCGTT ATCGTTATAA AGACGTATAT TTATAAAGAC GTATATTTAA CATTGACTAC ATTCTCCAAA
1121  CATATTGCTA ATAGCAGCTA CAATCCAGCT ACCATTCTGC TTTTATTTTA TGGTTGGGAT AAGGCTGGAT
      GTATAACGAT TATCGTCGAT GTTAGGTCGA TGGTAAGACG AAAATAAAAT ACCAACCCTA TTCCGACCTA
1191  TATTCTGAGT CCAAGCTAGG CCCTTTTGCT AATCATGTTC ATACCTCTTA TCTTCCTCCC ACAGCTCCTG
      ATAAGACTCA GGTTCGATCC GGGAAAACGA TTAGTACAAG TATGGAGAAT AGAAGGAGGG TGTCGAGGAC
1261  GGCAACGTGC TGGTCTGTGT GCTGGCCCAT CACTTTGGCA AAGAATTGGG ATTCGAACAT CGATTGAATT
      CCGTTGCACG ACCAGACACA CGACCGGGTA GTGAAACCGT TTCTTAACCC TAAGCTTGTA GCTAACTTAA
1331  CATGGAGTCT CCCTCGGCCC CTCCCCACAG ATGGTGCATC CCCTGGCAGA GGCTCCTGCT CACAGGTGAA
```

GTACCTCAGA GGGAGCCGGG GAGGGGTGTC TACCACGTAG GGGACCGTCT CCGAGGACGA GTGTCCACTT

1401 GGGAGGACAA CCTGGGAGAG GGTGGGAGGA GGGAGCTGGG GTCTCCTGGG TAGGACAGGG CTGTCTAGAA

CCCTCCTGTT GGACCCTCTC CCACCCTCCT CCCTCGACCC CAGAGGACCC ATCCTGTCCC GACAGATCTT

1471 AAGTCGAGGC GGAGGTACAA ATTGACAGGT TAATTACAGG CAGACTTCAA AGCCTTCAAA CCTATGTAAC

TTCAGCTCCG CCTCCATGTT TAACTGTCCA ATTAATGTCC GTCTGAAGTT TCGGAAGTTT GGATACATTG

1541 ACAACAACTA ATCAGGATTA AATGGCCTTG GTATGTTCCA CTCGGCACCA TTATTGGACT AATTGCCATC

TGTTGTTGAT TAGTCCTAAT TTACCGGAAC CATACAAGGT GAGCCGTGGT AATAACCTGA TTAACGGTAG

1611 GTCATGGTTA CAATCTTGCT TTGTTGCATG ACTAAGCTTG CCTCGAGCAG CGCTGCTCGA GAGATCTACG

CAGTACCAAT GTTAGAACGA AACAACGTAC TGATTCGAAC GGAGCTCGTC GCGACGAGCT CTCTAGATGC

1681 GGTGGCATCC CTGTGACCCC TCCCCAGTGC CTCTCCTGGC CCTGGAAGTT GCCACTCCAG TGCCCACCAG

CCACCGTAGG GACACTGGGG AGGGGTCACG GAGAGGACCG GGACCTTCAA CGGTGAGGTC ACGGGTGGTC

1751 CCTTGTCCTA ATAAAATTAA GTTGCATCAT TTTGTCTGAC TAGGTGTCCT TCTATAATAT TATGGGGTGG

GGAACAGGAT TATTTTAATT CAACGTAGTA AAACAGACTG ATCCACAGGA AGATATTATA ATACCCCACC

1821 AGGGGGGTGG TATGGAGCAA GGGGCAAGTT GGGAAGACAA CCTGTAGGGC CTGCGGGGTC TATTGGGAAC

TCCCCCCACC ATACCTCGTT CCCCGTTCAA CCCTTCTGTT GGACATCCCG GACGCCCCAG ATAACCCTTG

1891 CAAGCTGGAG TGCAGTGGCA CAATCTTGGC TCACTGCAAT CTCCGCCTCC TGGGTTCAAG CGATTCTCCT

GTTCGACCTC ACGTCACCGT GTTAGAACCG AGTGACGTTA GAGGCGGAGG ACCCAAGTTC GCTAAGAGGA

1961 GCCTCAGCCT CCCGAGTTGT TGGGATTCCA GGCATGCATG ACCAGGCTCA GCTAATTTTT GTTTTTTTGG

CGGAGTCGGA GGGCTCAACA ACCCTAAGGT CCGTACGTAC TGGTCCGAGT CGATTAAAAA CAAAAAAACC

2031 TAGAGACGGG GTTTCACCAT ATTGGCCAGG CTGGTCTCCA ACTCCTAATC TCAGGTGATC TACCCACCTT

ATCTCTGCCC CAAAGTGGTA TAACCGGTCC GACCAGAGGT TGAGGATTAG AGTCCACTAG ATGGGTGGAA

2101 GGCCTCCCAA ATTGCTGGGA TTACAGGCGT GAACCACTGC TCCCTTCCCT GTCCTTCTGA TTTTGTAGGT

CCGGAGGGTT TAACGACCCT AATGTCCGCA CTTGGTGACG AGGGAAGGGA CAGGAAGACT AAAACATCCA

2171 AACCACGTGC GGACCGAGCG GCCGCAGGAA CCCCTAGTGA TGGAGTTGGC CACTCCCTCT CTGCGCGCTC

TTGGTGCACG CCTGGCTCGC CGGCGTCCTT GGGGATCACT ACCTCAACCG GTGAGGGAGA GACGCGCGAG

2241 GCTCGCTCAC TGAGGCCGGG CGACCAAAGG TCGCCCGACG CCCGGGCTTT GCCCGGGCGG CCTCAGTGAG

CGAGCGAGTG ACTCCGGCCC GCTGGTTTCC AGCGGGCTGC GGGCCCGAAA CGGGCCCGCC GGAGTCACTC

2311 CGAGCGAGCG CGCAGCTGCC TGCAGGGGCG CCTGATGCGG TATTTTCTCC TTACGCATCT GTGCGGTATT

GCTCGCTCGC GCGTCGACGG ACGTCCCCGC GGACTACGCC ATAAAAGAGG AATGCGTAGA CACGCCATAA

2381 TCACACCGCA TACGTCAAAG CAACCATAGT ACGCGCCCTG TAGCGGCGCA TTAAGCGCGG CGGGTGTGGT

AGTGTGGCGT ATGCAGTTTC GTTGGTATCA TGCGCGGGAC ATCGCCGCGT AATTCGCGCC GCCCACACCA

2451 GGTTACGCGC AGCGTGACCG CTACACTTGC CAGCGCCCTA GCGCCCGCTC CTTTCGCTTT CTTCCCTTCC

CCAATGCGCG TCGCACTGGC GATGTGAACG GTCGCGGGAT CGCGGGCGAG GAAAGCGAAA GAAGGGAAGG

2521 TTTCTCGCCA CGTTCGCCGG CTTTCCCCGT CAAGCTCTAA ATCGGGGGCT CCCTTTAGGG TTCCGATTTA

AAAGAGCGGT GCAAGCGGCC GAAAGGGGCA GTTCGAGATT TAGCCCCCGA GGGAAATCCC AAGGCTAAAT

2591 GTGCTTTACG GCACCTCGAC CCCAAAAAAC TTGATTTGGG TGATGGTTCA CGTAGTGGGC CATCGCCCTG

CACGAAATGC CGTGGAGCTG GGGTTTTTTG AACTAAACCC ACTACCAAGT GCATCACCCG GTAGCGGGAC

2661 ATAGACGGTT TTTCGCCCTT TGACGTTGGA GTCCACGTTC TTTAATAGTG GACTCTTGTT CCAAACTGGA

TATCTGCCAA AAAGCGGGAA ACTGCAACCT CAGGTGCAAG AAATTATCAC CTGAGAACAA GGTTTGACCT

2731 ACAACACTCA ACCCTATCTC GGGCTATTCT TTTGATTTAT AAGGGATTTT GCCGATTTCG GCCTATTGGT

TGTTGTGAGT TGGGATAGAG CCCGATAAGA AAACTAAATA TTCCCTAAAA CGGCTAAAGC CGGATAACCA

2801 TAAAAAATGA GCTGATTTAA CAAAAATTTA ACGCGAATTT TAACAAAATA TTAACGTTTA CAATTTTATG

ATTTTTTACT CGACTAAATT GTTTTTAAAT TGCGCTTAAA ATTGTTTTAT AATTGCAAAT GTTAAAATAC

2871 GTGCACTCTC AGTACAATCT GCTCTGATGC CGCATAGTTA AGCCAGCCCC GACACCCGCC AACACCCGCT

CACGTGAGAG TCATGTTAGA CGAGACTACG GCGTATCAAT TCGGTCGGGG CTGTGGGCGG TTGTGGGCGA

2941 GACGCGCCCT GACGGGCTTG TCTGCTCCCG GCATCCGCTT ACAGACAAGC TGTGACCGTC TCCGGGAGCT

CTGCGCGGGA CTGCCCGAAC AGACGAGGGC CGTAGGCGAA TGTCTGTTCG ACACTGGCAG AGGCCCTCGA

3011 GCATGTGTCA GAGGTTTTCA CCGTCATCAC CGAAACGCGC GAGACGAAAG GGCCTCGTGA TACGCCTATT

CGTACACAGT CTCCAAAAGT GGCAGTAGTG GCTTTGCGCG CTCTGCTTTC CCGGAGCACT ATGCGGATAA

3081 TTTATAGGTT AATGTCATGA TAATAATGGT TTCTTAGACG TCAGGTGGCA CTTTTCGGGG AAATGTGCGC

AAATATCCAA TTACAGTACT ATTATTACCA AAGAATCTGC AGTCCACCGT GAAAAGCCCC TTTACACGCG

3151 GGAACCCCTA TTTGTTTATT TTTCTAAATA CATTCAAATA TGTATCCGCT CATGAGACAA TAACCCTGAT

CCTTGGGGAT AAACAAATAA AAAGATTTAT GTAAGTTTAT ACATAGGCGA GTACTCTGTT ATTGGGACTA

3221 AAATGCTTCA ATAATATTGA AAAAGGAAGA GTATGAGTAT TCAACATTTC CGTGTCGCCC TTATTCCCTT

TTTACGAAGT TATTATAACT TTTTCCTTCT CATACTCATA AGTTGTAAAG GCACAGCGGG AATAAGGGAA

3291 TTTTGCGGCA TTTTGCCTTC CTGTTTTTGC TCACCCAGAA ACGCTGGTGA AAGTAAAAGA TGCTGAAGAT

AAAACGCCGT AAAACGGAAG GACAAAAACG AGTGGGTCTT TGCGACCACT TTCATTTTCT ACGACTTCTA

3361 CAGTTGGGTG CACGAGTGGG TTACATCGAA CTGGATCTCA ACAGCGGTAA GATCCTTGAG AGTTTTCGCC

GTCAACCCAC GTGCTCACCC AATGTAGCTT GACCTAGAGT TGTCGCCATT CTAGGAACTC TCAAAAGCGG

3431 CCGAAGAACG TTTTCCAATG ATGAGCACTT TTAAAGTTCT GCTATGTGGC GCGGTATTAT CCCGTATTGA

GGCTTCTTGC AAAAGGTTAC TACTCGTGAA AATTTCAAGA CGATACACCG CGCCATAATA GGGCATAACT

3501 CGCCGGGCAA GAGCAACTCG GTCGCCGCAT ACACTATTCT CAGAATGACT TGGTTGAGTA CTCACCAGTC

GCGGCCCGTT CTCGTTGAGC CAGCGGCGTA TGTGATAAGA GTCTTACTGA ACCAACTCAT GAGTGGTCAG

3571 ACAGAAAAGC ATCTTACGGA TGGCATGACA GTAAGAGAAT TATGCAGTGC TGCCATAACC ATGAGTGATA

TGTCTTTTCG TAGAATGCCT ACCGTACTGT CATTCTCTTA ATACGTCACG ACGGTATTGG TACTCACTAT

3641 ACACTGCGGC CAACTTACTT CTGACAACGA TCGGAGGACC GAAGGAGCTA ACCGCTTTTT TGCACAACAT

TGTGACGCCG GTTGAATGAA GACTGTTGCT AGCCTCCTGG CTTCCTCGAT TGGCGAAAAA ACGTGTTGTA

3711 GGGGGATCAT GTAACTCGCC TTGATCGTTG GGAACCGGAG CTGAATGAAG CCATACCAAA CGACGAGCGT

CCCCCTAGTA CATTGAGCGG AACTAGCAAC CCTTGGCCTC GACTTACTTC GGTATGGTTT GCTGCTCGCA

3781 GACACCACGA TGCCTGTAGC AATGGCAACA ACGTTGCGCA AACTATTAAC TGGCGAACTA CTTACTCTAG

CTGTGGTGCT ACGGACATCG TTACCGTTGT TGCAACGCGT TTGATAATTG ACCGCTTGAT GAATGAGATC

3851 CTTCCCGGCA ACAATTAATA GACTGGATGG AGGCGGATAA AGTTGCAGGA CCACTTCTGC GCTCGGCCCT

GAAGGGCCGT TGTTAATTAT CTGACCTACC TCCGCCTATT TCAACGTCCT GGTGAAGACG CGAGCCGGGA

3921 TCCGGCTGGC TGGTTTATTG CTGATAAATC TGGAGCCGGT GAGCGTGGGT CTCGCGGTAT CATTGCAGCA

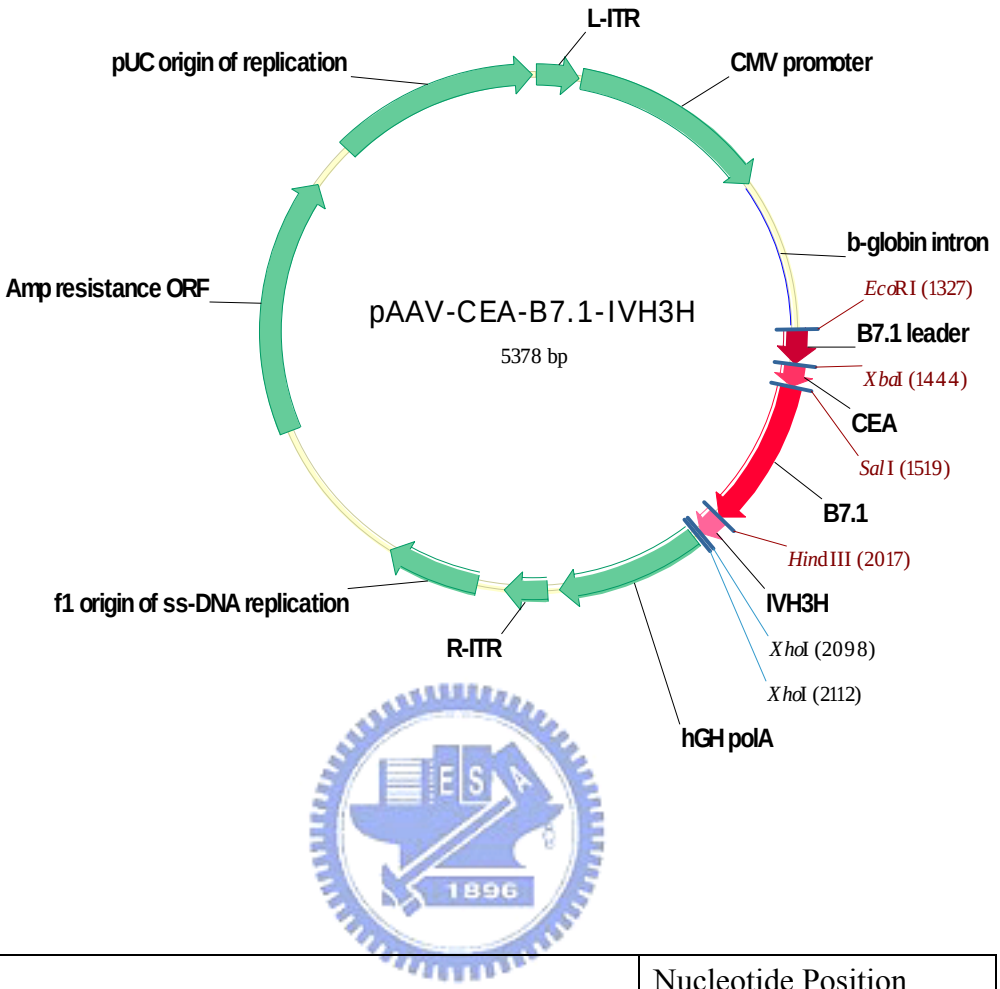
AGGCCGACCG ACCAAATAAC GACTATTTAG ACCTCGGCCA CTCGCACCCA GAGCGCCATA GTAACGTCGT

3991 CTGGGGCCAG ATGGTAAGCC CTCCCGTATC GTAGTTATCT ACACGACGGG GAGTCAGGCA ACTATGGATG

GACCCCGGTC TACCATTCGG GAGGGCATAG CATCAATAGA TGTGCTGCCC CTCAGTCCGT TGATACCTAC

4061 AACGAAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT TAAGCATTGG TAACTGTCAG ACCAAGTTTA

```
       TTGCTTTATC TGTCTAGCGA CTCTATCCAC GGAGTGACTA ATTCGTAACC ATTGACAGTC TGGTTCAAAT
4131   CTCATATATA CTTTAGATTG ATTTAAAACT TCATTTTTAA TTTAAAAGGA TCTAGGTGAA GATCCTTTTT
       GAGTATATAT GAAATCTAAC TAAATTTTGA AGTAAAAATT AAATTTTCCT AGATCCACTT CTAGGAAAAA
4201   GATAATCTCA TGACCAAAAT CCCTTAACGT GAGTTTTCGT TCCACTGAGC GTCAGACCCC GTAGAAAAGA
       CTATTAGAGT ACTGGTTTTA GGGAATTGCA CTCAAAAGCA AGGTGACTCG CAGTCTGGGG CATCTTTTCT
4271   TCAAAGGATC TTCTTGAGAT CCTTTTTTTC TGCGCGTAAT CTGCTGCTTG CAAACAAAAA AACCACCGCT
       AGTTTCCTAG AAGAACTCTA GGAAAAAAAG ACGCGCATTA GACGACGAAC GTTTGTTTTT TTGGTGGCGA
4341   ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA AGGTAACTGG CTTCAGCAGA
       TGGTCGCCAC CAAACAAACG GCCTAGTTCT CGATGGTTGA GAAAAAGGCT TCCATTGACC GAAGTCGTCT
4411   GCGCAGATAC CAAATACTGT CCTTCTAGTG TAGCCGTAGT TAGGCCACCA CTTCAAGAAC TCTGTAGCAC
       CGCGTCTATG GTTTATGACA GGAAGATCAC ATCGGCATCA ATCCGGTGGT GAAGTTCTTG AGACATCGTG
4481   CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAGTGGC TGCTGCCAGT GGCGATAAGT CGTGTCTTAC
       GCGGATGTAT GGAGCGAGAC GATTAGGACA ATGGTCACCG ACGACGGTCA CCGCTATTCA GCACAGAATG
4551   CGGGTTGGAC TCAAGACGAT AGTTACCGGA TAAGGCGCAG CGGTCGGGCT GAACGGGGGG TTCGTGCACA
       GCCCAACCTG AGTTCTGCTA TCAATGGCCT ATTCCGCGTC GCCAGCCCGA CTTGCCCCCC AAGCACGTGT
4621   CAGCCCAGCT TGGAGCGAAC GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCTATGA GAAAGCGCCA
       GTCGGGTCGA ACCTCGCTTG CTGGATGTGG CTTGACTCTA TGGATGTCGC ACTCGATACT CTTTCGCGGT
4691   CGCTTCCCGA AGGGAGAAAG GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG AGCGCACGAG
       GCGAAGGGCT TCCCTCTTTC CGCCTGTCCA TAGGCCATTC GCCGTCCCAG CCTTGTCCTC TCGCGTGCTC
4761   GGAGCTTCCA GGGGGAAACG CCTGGTATCT TTATAGTCCT GTCGGGTTTC GCCACCTCTG ACTTGAGCGT
       CCTCGAAGGT CCCCCTTTGC GGACCATAGA AATATCAGGA CAGCCCAAAG CGGTGGAGAC TGAACTCGCA
4831   CGATTTTTGT GATGCTCGTC AGGGGGGCGG AGCCTATGGA AAAACGCCAG CAACGCGGCC TTTTTACGGT
       GCTAAAAACA CTACGAGCAG TCCCCCCGCC TCGGATACCT TTTTGCGGTC GTTGCGCCGG AAAAATGCCA
4901   TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGT
       AGGACCGGAA AACGACCGGA AAACGAGTGT ACA
```



Feature	Nucleotide Position
Left AAV-2 inverted terminal repeat (ITR)	1-141
CMV promoter	150-812
β-globin intron	820-1312
B7.1 leader	1332-1442
CEA	1448-1517
B7.1	1524-2015
IVH3H	2022-2096
Human growth hormone (hGH) polyA signal	2123-2601
Right AAV-2 inverted terminal repeat (ITR)	2641-2781
fl origin of ss-DNA replication	2873-3179
Ampicillin resistance (*bla*) ORF	3698-4555
pUC origin of replication	4706-5373

```
   1  CCTGCAGGCA GCTGCGCGCT CGCTCGCTCA CTGAGGCCGC CCGGGCAAAG CCCGGGCGTC GGGCGACCTT
      GGACGTCCGT CGACGCGCGA GCGAGCGAGT GACTCCGGCG GGCCCGTTTC GGGCCCGCAG CCCGCTGGAA
  71  TGGTCGCCCG GCCTCAGTGA GCGAGCGAGC GCGCAGAGAG GGAGTGGCCA ACTCCATCAC TAGGGGTTCC
      ACCAGCGGGC CGGAGTCACT CGCTCGCTCG CGCGTCTCTC CCTCACCGGT TGAGGTAGTG ATCCCCAAGG
 141  TGCGGCCGCA CGCGTGGAGC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
      ACGCCGGCGT GCGCACCTCG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT
 211  TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCCGCCCATT
      ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA
 281  GACGTCAATA ATGACGTATG TTCCCATAGT AACGTCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG
      CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCAGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC
 351  TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC CCTATTGACG
      ATAAATGCCA TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCATGCGGG GGATAACTGC
 421  TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA
      AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT ACCCTGAAAG GATGAACCGT
 491  GTACATCTAC GTATTAGTCA TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA
      CATGTAGATG CATAATCAGT AGCGATAATG GTACCACTAC GCCAAAACCG TCATGTAGTT ACCCGCACCT
 561  TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGCACCA
      ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC AAAACGTGGT
 631  AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC AAATGGGCGG TAGGCGTGTA
      TTTAGTTGCC CTGAAAGGTT TTACAGCATT GTTGAGGCGG GGTAACTGCG TTTACCCGCC ATCCGCACAT
 701  CGGTGGGAGG TCTATATAAG CAGAGCTCGT TTAGTGAACC GTCAGATCGC CTGGAGACGC CATCCACGCT
      GCCACCCTCC AGATATATTC GTCTCGAGCA AATCACTTGG CAGTCTAGCG GACCTCTGCG GTAGGTGCGA
 771  GTTTTGACCT CCATAGAAGA CACCGGGACC GATCCAGCCT CCGCGGATTC GAATCCCGGC CGGGAACGGT
      CAAAACTGGA GGTATCTTCT GTGGCCCTGG CTAGGTCGGA GGCGCCTAAG CTTAGGGCCG GCCCTTGCCA
 841  GCATTGGAAC GCGGATTCCC CGTGCCAAGA GTGACGTAAG TACCGCCTAT AGAGTCTATA GGCCCACAAA
      CGTAACCTTG CGCCTAAGGG GCACGGTTCT CACTGCATTC ATGGCGGATA TCTCAGATAT CCGGGTGTTT
 911  AAATGCTTTC TTCTTTTAAT ATACTTTTTT GTTTATCTTA TTTCTAATAC TTTCCCTAAT CTCTTTCTTT
      TTTACGAAAG AAGAAAATTA TATGAAAAAA CAAATAGAAT AAAGATTATG AAAGGGATTA GAGAAAGAAA
 981  CAGGGCAATA ATGATACAAT GTATCATGCC TCTTTGCACC ATTCTAAAGA ATAACAGTGA TAATTTCTGG
      GTCCCGTTAT TACTATGTTA CATAGTACGG AGAAACGTGG TAAGATTTCT TATTGTCACT ATTAAAGACC
1051  GTTAAGGCAA TAGCAATATT TCTGCATATA AATATTTCTG CATATAAATT GTAACTGATG TAAGAGGTTT
      CAATTCCGTT ATCGTTATAA AGACGTATAT TTATAAAGAC GTATATTTAA CATTGACTAC ATTCTCCAAA
1121  CATATTGCTA ATAGCAGCTA CAATCCAGCT ACCATTCTGC TTTTATTTTA TGGTTGGGAT AAGGCTGGAT
      GTATAACGAT TATCGTCGAT GTTAGGTCGA TGGTAAGACG AAAATAAAAT ACCAACCCTA TTCCGACCTA
1191  TATTCTGAGT CCAAGCTAGG CCCTTTTGCT AATCATGTTC ATACCTCTTA TCTTCCTCCC ACAGCTCCTG
      ATAAGACTCA GGTTCGATCC GGGAAAACGA TTAGTACAAG TATGGAGAAT AGAAGGAGGG TGTCGAGGAC
1261  GGCAACGTGC TGGTCTGTGT GCTGGCCCAT CACTTTGGCA AAGAATTGGG ATTCGAACAT CGATTGAATT
      CCGTTGCACG ACCAGACACA CGACCGGGTA GTGAAACCGT TTCTTAACCC TAAGCTTGTA GCTAACTTAA
1331  CATGGCTTGC AATTGTCAGT TGATGCAGGA TACACCACTC CTCAAGTTTC CATGTCCAAG GCTCATTCTT
```

GTACCGAACG TTAACAGTCA ACTACGTCCT ATGTGGTGAG GAGTTCAAAG GTACAGGTTC CGAGTAAGAA

1401 CTCTTTGTGC TGCTGATTCG TCTTTCACAA GTGTCTTCAG ATTCTAGAGG TGAAGGGAGG ACAACCTGGG

GAGAAACACG ACGACTAAGC AGAAAGTGTT CACAGAAGTC TAAGATCTCC ACTTCCCTCC TGTTGGACCC

1471 AGAGGGTGGG AGGAGGGAGC TGGGGTCTCC TGGGTAGGAC AGGGCTGGTC GACAACGGAT CTTGCAATGG

TCTCCCACCC TCCTCCCTCG ACCCCAGAGG ACCCATCCTG TCCCGACCAG CTGTTGCCTA GAACGTTACC

1541 CAGCTGTTGC AATGGCAGCT GTTGCAACGG ATCTTGTAAG CTTTGCCCAA AGTACGTGAA GCAAAACACA

GTCGACAACG TTACCGTCGA CAACGTTGCC TAGAACATTC GAAACGGGTT TCATGCACTT CGTTTTGTGT

1611 CTTAAACTGG CTACCGGAAT GAGAAACGTG CCAGAAAAGC AAACATAACT CGAGCAGCGC TGCTCGAGAG

GAATTTGACC GATGGCCTTA CTCTTTGCAC GGTCTTTTCG TTTGTATTGA GCTCGTCGCG ACGAGCTCTC

1681 ATCTACGGGT GGCATCCCTG TGACCCCTCC CCAGTGCCTC TCCTGGCCCT GGAAGTTGCC ACTCCAGTGC

TAGATGCCCA CCGTAGGGAC ACTGGGGAGG GGTCACGGAG AGGACCGGGA CCTTCAACGG TGAGGTCACG

1751 CCACCAGCCT TGTCCTAATA AAATTAAGTT GCATCATTTT GTCTGACTAG GTGTCCTTCT ATAATATTAT

GGTGGTCGGA ACAGGATTAT TTTAATTCAA CGTAGTAAAA CAGACTGATC CACAGGAAGA TATTATAATA

1821 GGGGTGGAGG GGGGTGGTAT GGAGCAAGGG GCAAGTTGGG AAGACAACCT GTAGGGCCTG CGGGGTCTAT

CCCCACCTCC CCCCACCATA CCTCGTTCCC CGTTCAACCC TTCTGTTGGA CATCCCGGAC GCCCCAGATA

1891 TGGGAACCAA GCTGGAGTGC AGTGGCACAA TCTTGGCTCA CTGCAATCTC CGCCTCCTGG GTTCAAGCGA

ACCCTTGGTT CGACCTCACG TCACCGTGTT AGAACCGAGT GACGTTAGAG GCGGAGGACC CAAGTTCGCT

1961 TTCTCCTGCC TCAGCCTCCC GAGTTGTTGG GATTCCAGGC ATGCATGACC AGGCTCAGCT AATTTTTGTT

AAGAGGACGG AGTCGGAGGG CTCAACAACC CTAAGGTCCG TACGTACTGG TCCGAGTCGA TTAAAAACAA

2031 TTTTTGGTAG AGACGGGGTT TCACCATATT GGCCAGGCTG GTCTCCAACT CCTAATCTCA GGTGATCTAC

AAAAACCATC TCTGCCCCAA AGTGGTATAA CCGGTCCGAC CAGAGGTTGA GGATTAGAGT CCACTAGATG

2101 CCACCTTGGC CTCCCAAATT GCTGGGATTA CAGGCGTGAA CCACTGCTCC CTTCCCTGTC CTTCTGATTT

GGTGGAACCG GAGGGTTTAA CGACCCTAAT GTCCGCACTT GGTGACGAGG GAAGGGACAG GAAGACTAAA

2171 TGTAGGTAAC CACGTGCGGA CCGAGCGGCC GCAGGAACCC CTAGTGATGG AGTTGGCCAC TCCCTCTCTG

ACATCCATTG GTGCACGCCT GGCTCGCCGG CGTCCTTGGG GATCACTACC TCAACCGGTG AGGGAGAGAC

2241 CGCGCTCGCT CGCTCACTGA GGCCGGGCGA CCAAAGGTCG CCCGACGCCC GGGCTTTGCC CGGGCGGCCT

GCGCGAGCGA GCGAGTGACT CCGGCCCGCT GGTTTCCAGC GGGCTGCGGG CCCGAAACGG GCCCGCCGGA

2311 CAGTGAGCGA GCGAGCGCGC AGCTGCCTGC AGGGGCGCCT GATGCGGTAT TTTCTCCTTA CGCATCTGTG

GTCACTCGCT CGCTCGCGCG TCGACGGACG TCCCCGCGGA CTACGCCATA AAAGAGGAAT GCGTAGACAC

2381 CGGTATTTCA CACCGCATAC GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATTA AGCGCGGCGG

GCCATAAAGT GTGGCGTATG CAGTTTCGTT GGTATCATGC GCGGGACATC GCCGCGTAAT TCGCGCCGCC

2451 GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT TCGCTTTCTT

CACACCACCA ATGCGCGTCG CACTGGCGAT GTGAACGGTC GCGGGATCGC GGGCGAGGAA AGCGAAAGAA

2521 CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC GGGGGCTCCC TTTAGGGTTC

GGGAAGGAAA GAGCGGTGCA AGCGGCCGAA AGGGGCAGTT CGAGATTTAG CCCCCGAGGG AAATCCCAAG

2591 CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAAACTTG ATTTGGGTGA TGGTTCACGT AGTGGGCCAT

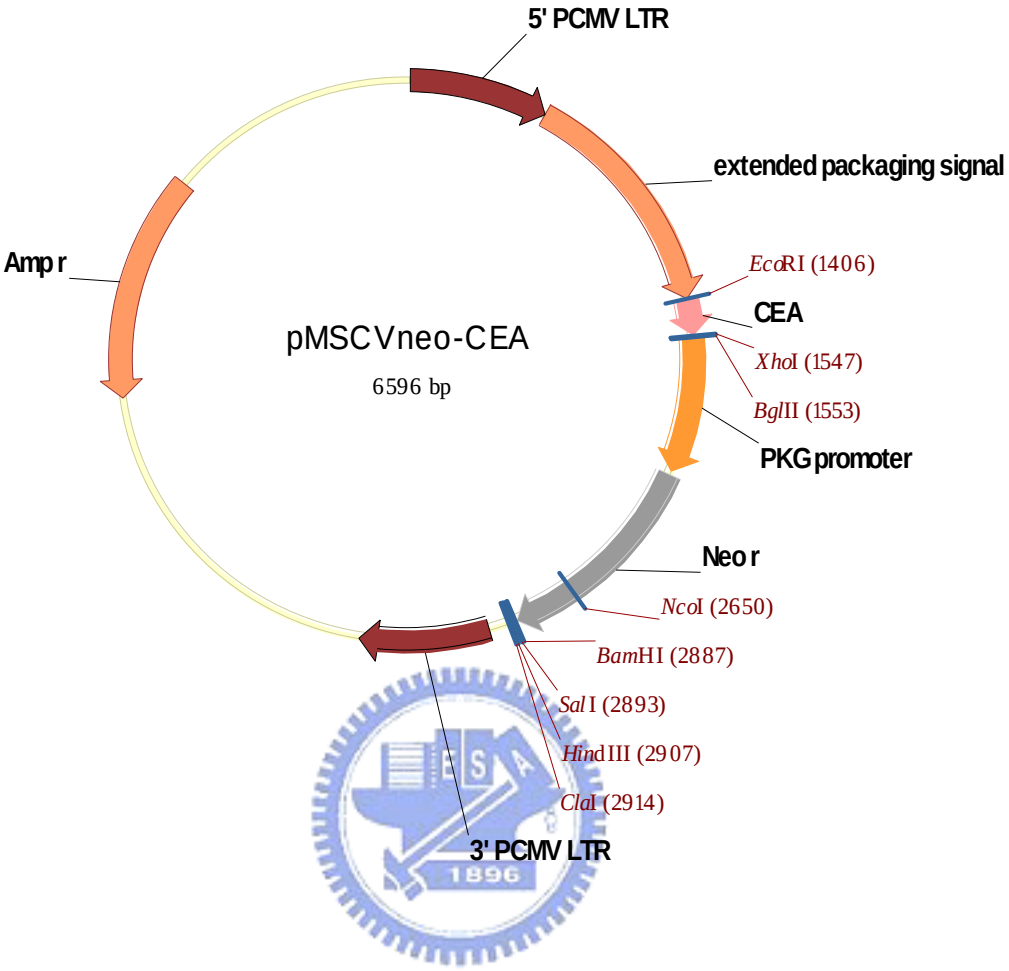
GCTAAATCAC GAAATGCCGT GGAGCTGGGG TTTTTTGAAC TAAACCCACT ACCAAGTGCA TCACCCGGTA

2661 CGCCCTGATA GACGGTTTTT CGCCCTTTGA CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTTCCA

GCGGGACTAT CTGCCAAAAA GCGGGAAACT GCAACCTCAG GTGCAAGAAA TTATCACCTG AGAACAAGGT

```
2731  AACTGGAACA ACACTCAACC CTATCTCGGG CTATTCTTTT GATTTATAAG GGATTTTGCC GATTTCGGCC
      TTGACCTTGT TGTGAGTTGG GATAGAGCCC GATAAGAAAA CTAAATATTC CCTAAAACGG CTAAAGCCGG
2801  TATTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAACG CGAATTTTAA CAAAATATTA ACGTTTACAA
      ATAACCAATT TTTTACTCGA CTAAATTGTT TTTAAATTGC GCTTAAAATT GTTTTATAAT TGCAAATGTT
2871  TTTTATGGTG CACTCTCAGT ACAATCTGCT CTGATGCCGC ATAGTTAAGC CAGCCCCGAC ACCCGCCAAC
      AAAATACCAC GTGAGAGTCA TGTTAGACGA GACTACGGCG TATCAATTCG GTCGGGGCTG TGGGCGGTTG
2941  ACCCGCTGAC GCGCCCTGAC GGGCTTGTCT GCTCCCGGCA TCCGCTTACA GACAAGCTGT GACCGTCTCC
      TGGGCGACTG CGCGGGACTG CCCGAACAGA CGAGGGCCGT AGGCGAATGT CTGTTCGACA CTGGCAGAGG
3011  GGGAGCTGCA TGTGTCAGAG GTTTTCACCG TCATCACCGA AACGCGCGAG ACGAAAGGGC CTCGTGATAC
      CCCTCGACGT ACACAGTCTC CAAAAGTGGC AGTAGTGGCT TTGCGCGCTC TGCTTTCCCG GAGCACTATG
3081  GCCTATTTTT ATAGGTTAAT GTCATGATAA TAATGGTTTC TTAGACGTCA GGTGGCACTT TTCGGGGAAA
      CGGATAAAAA TATCCAATTA CAGTACTATT ATTACCAAAG AATCTGCAGT CCACCGTGAA AAGCCCCTTT
3151  TGTGCGCGGA ACCCCTATTT GTTTATTTTT CTAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA
      ACACGCGCCT TGGGGATAAA CAAATAAAAA GATTTATGTA AGTTTATACA TAGGCGAGTA CTCTGTTATT
3221  CCCTGATAAA TGCTTCAATA ATATTGAAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT GTCGCCCTTA
      GGGACTATTT ACGAAGTTAT TATAACTTTT TCCTTCTCAT ACTCATAAGT TGTAAAGGCA CAGCGGGAAT
3291  TTCCCTTTTT TGCGGCATTT TGCCTTCCTG TTTTTGCTCA CCCAGAAACG CTGGTGAAAG TAAAAGATGC
      AAGGGAAAAA ACGCCGTAAA ACGGAAGGAC AAAAACGAGT GGGTCTTTGC GACCACTTTC ATTTTCTACG
3361  TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT CCTTGAGAGT
      ACTTCTAGTC AACCCACGTG CTCACCCAAT GTAGCTTGAC CTAGAGTTGT CGCCATTCTA GGAACTCTCA
3431  TTTCGCCCCG AAGAACGTTT TCCAATGATG AGCACTTTTA AAGTTCTGCT ATGTGGCGCG GTATTATCCC
      AAAGCGGGGC TTCTTGCAAA AGGTTACTAC TCGTGAAAAT TTCAAGACGA TACACCGCGC CATAATAGGG
3501  GTATTGACGC CGGGCAAGAG CAACTCGGTC GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC
      CATAACTGCG GCCCGTTCTC GTTGAGCCAG CGGCGTATGT GATAAGAGTC TTACTGAACC AACTCATGAG
3571  ACCAGTCACA GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAACCATG
      TGGTCAGTGT CTTTTCGTAG AATGCCTACC GTACTGTCAT TCTCTTAATA CGTCACGACG GTATTGGTAC
3641  AGTGATAACA CTGCGGCCAA CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC GCTTTTTTGC
      TCACTATTGT GACGCCGGTT GAATGAAGAC TGTTGCTAGC CTCCTGGCTT CCTCGATTGG CGAAAAAACG
3711  ACAACATGGG GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA TACCAAACGA
      TGTTGTACCC CCTAGTACAT TGAGCGGAAC TAGCAACCCT TGGCCTCGAC TTACTTCGGT ATGGTTTGCT
3781  CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG TTGCGCAAAC TATTAACTGG CGAACTACTT
      GCTCGCACTG TGGTGCTACG GACATCGTTA CCGTTGTTGC AACGCGTTTG ATAATTGACC GCTTGATGAA
3851  ACTCTAGCTT CCCGGCAACA ATTAATAGAC TGGATGGAGG CGGATAAAGT TGCAGGACCA CTTCTGCGCT
      TGAGATCGAA GGGCCGTTGT TAATTATCTG ACCTACCTCC GCCTATTTCA ACGTCCTGGT GAAGACGCGA
3921  CGGCCCTTCC GGCTGGCTGG TTTATTGCTG ATAAATCTGG AGCCGGTGAG CGTGGGTCTC GCGGTATCAT
      GCCGGGAAGG CCGACCGACC AAATAACGAC TATTTAGACC TCGGCCACTC GCACCCAGAG CGCCATAGTA
3991  TGCAGCACTG GGGCCAGATG GTAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT
      ACGTCGTGAC CCCGGTCTAC CATTCGGGAG GGCATAGCAT CAATAGATGT GCTGCCCCTC AGTCCGTTGA
4061  ATGGATGAAC GAAATAGACA GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA CTGTCAGACC
```

```
      TACCTACTTG CTTTATCTGT CTAGCGACTC TATCCACGGA GTGACTAATT CGTAACCATT GACAGTCTGG
4131  AAGTTTACTC ATATATACTT TAGATTGATT TAAAACTTCA TTTTTAATTT AAAAGGATCT AGGTGAAGAT
      TTCAAATGAG TATATATGAA ATCTAACTAA ATTTTGAAGT AAAAATTAAA TTTTCCTAGA TCCACTTCTA
4201  CCTTTTTGAT AATCTCATGA CCAAAATCCC TTAACGTGAG TTTTCGTTCC ACTGAGCGTC AGACCCCGTA
      GGAAAAACTA TTAGAGTACT GGTTTTAGGG AATTGCACTC AAAAGCAAGG TGACTCGCAG TCTGGGGCAT
4271  GAAAAGATCA AAGGATCTTC TTGAGATCCT TTTTTTCTGC GCGTAATCTG CTGCTTGCAA ACAAAAAAAC
      CTTTTCTAGT TTCCTAGAAG AACTCTAGGA AAAAAAGACG CGCATTAGAC GACGAACGTT TGTTTTTTTG
4341  CACCGCTACC AGCGGTGGTT TGTTTGCCGG ATCAAGAGCT ACCAACTCTT TTTCCGAAGG TAACTGGCTT
      GTGGCGATGG TCGCCACCAA ACAAACGGCC TAGTTCTCGA TGGTTGAGAA AAAGGCTTCC ATTGACCGAA
4411  CAGCAGAGCG CAGATACCAA ATACTGTCCT TCTAGTGTAG CCGTAGTTAG GCCACCACTT CAAGAACTCT
      GTCGTCTCGC GTCTATGGTT TATGACAGGA AGATCACATC GGCATCAATC CGGTGGTGAA GTTCTTGAGA
4481  GTAGCACCGC CTACATACCT CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC TGCCAGTGGC GATAAGTCGT
      CATCGTGGCG GATGTATGGA GCGAGACGAT TAGGACAATG GTCACCGACG ACGGTCACCG CTATTCAGCA
4551  GTCTTACCGG GTTGGACTCA AGACGATAGT TACCGGATAA GGCGCAGCGG TCGGGCTGAA CGGGGGGTTC
      CAGAATGGCC CAACCTGAGT TCTGCTATCA ATGGCCTATT CCGCGTCGCC AGCCCGACTT GCCCCCCAAG
4621  GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA CTGAGATACC TACAGCGTGA GCTATGAGAA
      CACGTGTGTC GGGTCGAACC TCGCTTGCTG GATGTGGCTT GACTCTATGG ATGTCGCACT CGATACTCTT
4691  AGCGCCACGC TTCCCGAAGG GAGAAAGGCG GACAGGTATC CGGTAAGCGG CAGGGTCGGA ACAGGAGAGC
      TCGCGGTGCG AAGGGCTTCC CTCTTTCCGC CTGTCCATAG GCCATTCGCC GTCCCAGCCT TGTCCTCTCG
4761  GCACGAGGGA GCTTCCAGGG GGAAACGCCT GGTATCTTTA TAGTCCTGTC GGGTTTCGCC ACCTCTGACT
      CGTGCTCCCT CGAAGGTCCC CCTTTGCGGA CCATAGAAAT ATCAGGACAG CCCAAAGCGG TGGAGACTGA
4831  TGAGCGTCGA TTTTTGTGAT GCTCGTCAGG GGGGCGGAGC CTATGGAAAA ACGCCAGCAA CGCGGCCTTT
      ACTCGCAGCT AAAAACACTA CGAGCAGTCC CCCCGCCTCG GATACCTTTT TGCGGTCGTT GCGCCGGAAA
4901  TTACGGTTCC TGGCCTTTTG CTGGCCTTTT GCTCACATGT
      AATGCCAAGG ACCGGAAAAC GACCGGAAAA CGAGTGTACA
```



Feature	Nucleotide Position
5’ CMV promoter LTR	1-515
Extended packaging signal (Ψ)	516-1405
CEA	1411-1545
PKG promoter	1558-2066
3’ CMV promoter LTR	2988-3471
Ampicillin resistance (*bla*) ORF	5662-4805

```
   1  TGAAAGACCC CACCTGTAGG TTTGGCAAGC TAGCTTAAGT AACGCCATTT TGCAAGGCAT GGAAAATACA
      ACTTTCTGGG GTGGACATCC AAACCGTTCG ATCGAATTCA TTGCGGTAAA ACGTTCCGTA CCTTTTATGT
  71  TAACTGAGAA TAGAGAAGTT CAGATCAAGG TTAGGAACAG AGAGACAGCA GAATATGGGC CAAACAGGAT
      ATTGACTCTT ATCTCTTCAA GTCTAGTTCC AATCCTTGTC TCTCTGTCGT CTTATACCCG GTTTGTCCTA
 141  ATCTGTGGTA AGCAGTTCCT GCCCCGGCTC AGGGCCAAGA ACAGATGGTC CCCAGATGCG GTCCCGCCCT
      TAGACACCAT TCGTCAAGGA CGGGGCCGAG TCCCGGTTCT TGTCTACCAG GGGTCTACGC CAGGGCGGGA
 211  CAGCAGTTTC TAGAGAACCA TCAGATGTTT CCAGGGTGCC CCAAGGACCT GAAATGACCC TGTGCCTTAT
      GTCGTCAAAG ATCTCTTGGT AGTCTACAAA GGTCCCACGG GGTTCCTGGA CTTTACTGGG ACACGGAATA
 281  TTGAACTAAC CAATCAGTTC GCTTCTCGCT TCTGTTCGCG CGCTTCTGCT CCCCGAGCTC AATAAAAGAG
      AACTTGATTG GTTAGTCAAG CGAAGAGCGA AGACAAGCGC GCGAAGACGA GGGGCTCGAG TTATTTTCTC
 351  CCCACAACCC CTCACTCGGC GCGCCAGTCC TCCGATAGAC TGCGTCGCCC GGGTACCCGT ATTCCCAATA
      GGGTGTTGGG GAGTGAGCCG CGCGGTCAGG AGGCTATCTG ACGCAGCGGG CCCATGGGCA TAAGGGTTAT
 421  AAGCCTCTTG CTGTTTGCAT CCGAATCGTG GACTCGCTGA TCCTTGGGAG GGTCTCCTCA GATTGATTGA
      TTCGGAGAAC GACAAACGTA GGCTTAGCAC CTGAGCGACT AGGAACCCTC CCAGAGGAGT CTAACTAACT
 491  CTGCCCACCT CGGGGGTCTT TCATTTGGAG GTTCCACCGA GATTTGGAGA CCCCTGCCCA GGGACCACCG
      GACGGGTGGA GCCCCCAGAA AGTAAACCTC CAAGGTGGCT CTAAACCTCT GGGGACGGGT CCCTGGTGGC
 561  ACCCCCCCGC CGGGAGGTAA GCTGGCCAGC GGTCGTTTCG TGTCTGTCTC TGTCTTTGTG CGTGTTTGTG
      TGGGGGGGCG GCCCTCCATT CGACCGGTCG CCAGCAAAGC ACAGACAGAG ACAGAAACAC GCACAAACAC
 631  CCGGCATCTA ATGTTTGCGC CTGCGTCTGT ACTAGTTAGC TAACTAGCTC TGTATCTGGC GGACCCGTGG
      GGCCGTAGAT TACAAACGCG GACGCAGACA TGATCAATCG ATTGATCGAG ACATAGACCG CCTGGGCACC
 701  TGGAACTGAC GAGTTCTGAA CACCCGGCCG CAACCCTGGG AGACGTCCCA GGGACTTTGG GGGCCGTTTT
      ACCTTGACTG CTCAAGACTT GTGGGCCGGC GTTGGGACCC TCTGCAGGGT CCCTGAAACC CCCGGCAAAA
 771  TGTGGCCCGA CCTGAGGAAG GGAGTCGATG TGGAATCCGA CCCCGTCAGG ATATGTGGTT CTGGTAGGAG
      ACACCGGGCT GGACTCCTTC CCTCAGCTAC ACCTTAGGCT GGGGCAGTCC TATACACCAA GACCATCCTC
 841  ACGAGAACCT AAAACAGTTC CCGCCTCCGT CTGAATTTTT GCTTTCGGTT TGGAACCGAA GCCGCGCGTC
      TGCTCTTGGA TTTTGTCAAG GGCGGAGGCA GACTTAAAAA CGAAAGCCAA ACCTTGGCTT CGGCGCGCAG
 911  TTGTCTGCTG CAGCGCTGCA GCATCGTTCT GTGTTGTCTC TGTCTGACTG TGTTTCTGTA TTTGTCTGAA
      AACAGACGAC GTCGCGACGT CGTAGCAAGA CACAACAGAG ACAGACTGAC ACAAAGACAT AAACAGACTT
 981  AATTAGGGCC AGACTGTTAC CACTCCCTTA AGTTTGACCT TAGGTCACTG GAAAGATGTC GAGCGGATCG
      TTAATCCCGG TCTGACAATG GTGAGGGAAT TCAAACTGGA ATCCAGTGAC CTTTCTACAG CTCGCCTAGC
1051  CTCACAACCA GTCGGTAGAT GTCAAGAAGA GACGTTGGGT TACCTTCTGC TCTGCAGAAT GGCCAACCTT
      GAGTGTTGGT CAGCCATCTA CAGTTCTTCT CTGCAACCCA ATGGAAGACG AGACGTCTTA CCGGTTGGAA
1121  TAACGTCGGA TGGCCGCGAG ACGGCACCTT TAACCGAGAC CTCATCACCC AGGTTAAGAT CAAGGTCTTT
      ATTGCAGCCT ACCGGCGCTC TGCCGTGGAA ATTGGCTCTG GAGTAGTGGG TCCAATTCTA GTTCCAGAAA
1191  TCACCTGGCC CGCATGGACA CCCAGACCAG GTCCCCTACA TCGTGACCTG GGAAGCCTTG GCTTTTGACC
      AGTGGACCGG GCGTACCTGT GGGTCTGGTC CAGGGGATGT AGCACTGGAC CCTTCGGAAC CGAAAACTGG
1261  CCCCTCCCTG GGTCAAGCCC TTTGTACACC CTAAGCCTCC GCCTCCTCTT CCTCCATCCG CCCCGTCTCT
      GGGGAGGGAC CCAGTTCGGG AAACATGTGG GATTCGGAGG CGGAGGAGAA GGAGGTAGGC GGGGCAGAGA
1331  CCCCCTTGAA CCTCCTCGTT CGACCCCGCC TCGATCCTCC CTTTATCCAG CCCTCACTCC TTCTCTAGGC
```

GGGGGAACTT GGAGGAGCAA GCTGGGGCGG AGCTAGGAGG GAAATAGGTC GGGAGTGAGG AAGAGATCCG

1401 GCCGGAATTC ATGGAGTCTC CCTCGGCCCC TCCCCACAGA TGGTGCATCC CCTGGCAGAG GCTCCTGCTC

CGGCCTTAAG TACCTCAGAG GGAGCCGGGG AGGGGTGTCT ACCACGTAGG GGACCGTCTC CGAGGACGAG

1471 ACAGGTGAAG GGAGGACAAC CTGGGAGAGG GTGGGAGGAG GGAGCTGGGG TCTCCTGGGT AGGACAGGGC

TGTCCACTTC CCTCCTGTTG GACCCTCTCC CACCCTCCTC CCTCGACCCC AGAGGACCCA TCCTGTCCCG

1541 TGTAACTCGA GAGATCTAAT TCTACCGGGT AGGGGAGGCG CTTTTCCCAA GGCAGTCTGG AGCATGCGCT

ACATTGAGCT CTCTAGATTA AGATGGCCCA TCCCCTCCGC GAAAAGGGTT CCGTCAGACC TCGTACGCGA

1611 TTAGCAGCCC CGCTGGGCAC TTGGCGCTAC ACAAGTGGCC TCTGGCCTCG CACACATTCC ACATCCACCG

AATCGTCGGG GCGACCCGTG AACCGCGATG TGTTCACCGG AGACCGGAGC GTGTGTAAGG TGTAGGTGGC

1681 GTAGGCGCCA ACCGGCTCCG TTCTTTGGTG GCCCCTTCGC GCCACCTTCT ACTCCTCCCC TAGTCAGGAA

CATCCGCGGT TGGCCGAGGC AAGAAACCAC CGGGGAAGCG CGGTGGAAGA TGAGGAGGGG ATCAGTCCTT

1751 GTTCCCCCCC GCCCCGCAGC TCGCGTCGTG CAGGACGTGA CAAATGGAAG TAGCACGTCT CACTAGTCTC

CAAGGGGGGG CGGGGCGTCG AGCGCAGCAC GTCCTGCACT GTTTACCTTC ATCGTGCAGA GTGATCAGAG

1821 GTGCAGATGG ACAGCACCGC TGAGCAATGG AAGCGGGTAG GCCTTTGGGG CAGCGGCCAA TAGCAGCTTT

CACGTCTACC TGTCGTGGCG ACTCGTTACC TTCGCCCATC CGGAAACCCC GTCGCCGGTT ATCGTCGAAA

1891 GCTCCTTCGC TTTCTGGGCT CAGAGGCTGG GAAGGGGTGG GTCCGGGGGC GGGCTCAGGG GCGGGCTCAG

CGAGGAAGCG AAAGACCCGA GTCTCCGACC CTTCCCCACC CAGGCCCCCG CCCGAGTCCC CGCCCGAGTC

1961 GGGCGGGGCG GGCGCCCGAA GGTCCTCCGG AGGCCCGGCA TTCTGCACGC TTCAAAAGCG CACGTCTGCC

CCCGCCCCGC CCGCGGGCTT CCAGGAGGCC TCCGGGCCGT AAGACGTGCG AAGTTTTCGC GTGCAGACGG

2031 GCGCTGTTCT CCTCTTCCTC ATCTCCGGGC CTTTCGACCT GCAGCCAATA TGGGATCGGC CATTGAACAA

CGCGACAAGA GGAGAAGGAG TAGAGGCCCG GAAAGCTGGA CGTCGGTTAT ACCCTAGCCG GTAACTTGTT

2101 GATGGATTGC ACGCAGGTTC TCCGGCCGCT TGGGTGGAGA GGCTATTCGG CTATGACTGG GCACAACAGA

CTACCTAACG TGCGTCCAAG AGGCCGGCGA ACCCACCTCT CCGATAAGCC GATACTGACC CGTGTTGTCT

2171 CAATCGGCTG CTCTGATGCC GCCGTGTTCC GGCTGTCAGC GCAGGGGCGC CCGGTTCTTT TTGTCAAGAC

GTTAGCCGAC GAGACTACGG CGGCACAAGG CCGACAGTCG CGTCCCCGCG GGCCAAGAAA AACAGTTCTG

2241 CGACCTGTCC GGTGCCCTGA ATGAACTGCA GGACGAGGCA GCGCGGCTAT CGTGGCTGGC CACGACGGGC

GCTGGACAGG CCACGGGACT TACTTGACGT CCTGCTCCGT CGCGCCGATA GCACCGACCG GTGCTGCCCG

2311 GTTCCTTGCG CAGCTGTGCT CGACGTTGTC ACTGAAGCGG GAAGGGACTG GCTGCTATTG GGCGAAGTGC

CAAGGAACGC GTCGACACGA GCTGCAACAG TGACTTCGCC CTTCCCTGAC CGACGATAAC CCGCTTCACG

2381 CGGGGCAGGA TCTCCTGTCA TCTCACCTTG CTCCTGCCGA GAAAGTATCC ATCATGGCTG ATGCAATGCG

GCCCCGTCCT AGAGGACAGT AGAGTGGAAC GAGGACGGCT CTTTCATAGG TAGTACCGAC TACGTTACGC

2451 GCGGCTGCAT ACGCTTGATC CGGCTACCTG CCCATTCGAC CACCAAGCGA AACATCGCAT CGAGCGAGCA

CGCCGACGTA TGCGAACTAG GCCGATGGAC GGGTAAGCTG GTGGTTCGCT TTGTAGCGTA GCTCGCTCGT

2521 CGTACTCGGA TGGAAGCCGG TCTTGTCGAT CAGGATGATC TGGACGAAGA GCATCAGGGG CTCGCGCCAG

GCATGAGCCT ACCTTCGGCC AGAACAGCTA GTCCTACTAG ACCTGCTTCT CGTAGTCCCC GAGCGCGGTC

2591 CCGAACTGTT CGCCAGGCTC AAGGCGCGCA TGCCCGACGG CGAGGATCTC GTCGTGACCC ATGGCGATGC

GGCTTGACAA GCGGTCCGAG TTCCGCGCGT ACGGGCTGCC GCTCCTAGAG CAGCACTGGG TACCGCTACG

2661 CTGCTTGCCG AATATCATGG TGGAAAATGG CCGCTTTTCT GGATTCATCG ACTGTGGCCG GCTGGGTGTG

GACGAACGGC TTATAGTACC ACCTTTTACC GGCGAAAAGA CCTAAGTAGC TGACACCGGC CGACCCACAC

```
2731  GCGGACCGCT ATCAGGACAT AGCGTTGGCT ACCCGTGATA TTGCTGAAGA GCTTGGCGGC GAATGGGCTG
      CGCCTGGCGA TAGTCCTGTA TCGCAACCGA TGGGCACTAT AACGACTTCT CGAACCGCCG CTTACCCGAC
2801  ACCGCTTCCT CGTGCTTTAC GGTATCGCCG CTCCCGATTC GCAGCGCATC GCCTTCTATC GCCTTCTTGA
      TGGCGAAGGA GCACGAAATG CCATAGCGGC GAGGGCTAAG CGTCGCGTAG CGGAAGATAG CGGAAGAACT
2871  CGAGTTCTTC TGAGGGGATC CGTCGACCTG CAGCCAAGCT TATCGATAAA ATAAAAGATT TTATTTAGTC
      GCTCAAGAAG ACTCCCCTAG GCAGCTGGAC GTCGGTTCGA ATAGCTATTT TATTTTCTAA AATAAATCAG
2941  TCCAGAAAAA GGGGGGAATG AAAGACCCCA CCTGTAGGTT TGGCAAGCTA GCTTAAGTAA CGCCATTTTG
      AGGTCTTTTT CCCCCCTTAC TTTCTGGGGT GGACATCCAA ACCGTTCGAT CGAATTCATT GCGGTAAAAC
3011  CAAGGCATGG AAAATACATA ACTGAGAATA GAGAAGTTCA GATCAAGGTT AGGAACAGAG AGACAGCAGA
      GTTCCGTACC TTTTATGTAT TGACTCTTAT CTCTTCAAGT CTAGTTCCAA TCCTTGTCTC TCTGTCGTCT
3081  ATATGGGCCA AACAGGATAT CTGTGGTAAG CAGTTCCTGC CCCGGCTCAG GGCCAAGAAC AGATGGTCCC
      TATACCCGGT TTGTCCTATA GACACCATTC GTCAAGGACG GGGCCGAGTC CCGGTTCTTG TCTACCAGGG
3151  CAGATGCGGT CCCGCCCTCA GCAGTTTCTA GAGAACCATC AGATGTTTCC AGGGTGCCCC AAGGACCTGA
      GTCTACGCCA GGGCGGGAGT CGTCAAAGAT CTCTTGGTAG TCTACAAAGG TCCCACGGGG TTCCTGGACT
3221  AATGACCCTG TGCCTTATTT GAACTAACCA ATCAGTTCGC TTCTCGCTTC TGTTCGCGCG CTTCTGCTCC
      TTACTGGGAC ACGGAATAAA CTTGATTGGT TAGTCAAGCG AAGAGCGAAG ACAAGCGCGC GAAGACGAGG
3291  CCGAGCTCAA TAAAAGAGCC CACAACCCCT CACTCGGCGC GCCAGTCCTC CGATAGACTG CGTCGCCCGG
      GGCTCGAGTT ATTTTCTCGG GTGTTGGGGA GTGAGCCGCG CGGTCAGGAG GCTATCTGAC GCAGCGGGCC
3361  GTACCCGTGT ATCCAATAAA CCCTCTTGCA GTTGCATCCG ACTTGTGGTC TCGCTGTTCC TTGGGAGGGT
      CATGGGCACA TAGGTTATTT GGGAGAACGT CAACGTAGGC TGAACACCAG AGCGACAAGG AACCCTCCCA
3431  CTCCTCTGAG TGATTGACTA CCCGTCAGCG GGGGTCTTTC ATGGGTAACA GTTTCTTGAA GTTGGAGAAC
      GAGGAGACTC ACTAACTGAT GGGCAGTCGC CCCCAGAAAG TACCCATTGT CAAAGAACTT CAACCTCTTG
3501  AACATTCTGA GGGTAGGAGT CGAATATTAA GTAATCCTGA CTCAATTAGC CACTGTTTTG AATCCACATA
      TTGTAAGACT CCCATCCTCA GCTTATAATT CATTAGGACT GAGTTAATCG GTGACAAAAC TTAGGTGTAT
3571  CTCCAATACT CCTGAAATAG TTCATTATGG ACAGCGCAGA AGAGCTGGGG AGAATTAATT CGTAATCATG
      GAGGTTATGA GGACTTTATC AAGTAATACC TGTCGCGTCT TCTCGACCCC TCTTAATTAA GCATTAGTAC
3641  GTCATAGCTG TTTCCTGTGT GAAATTGTTA TCCGCTCACA ATTCCACACA ACATACGAGC CGGAAGCATA
      CAGTATCGAC AAAGGACACA CTTTAACAAT AGGCGAGTGT TAAGGTGTGT TGTATGCTCG GCCTTCGTAT
3711  AAGTGTAAAG CCTGGGGTGC CTAATGAGTG AGCTAACTCA CATTAATTGC GTTGCGCTCA CTGCCCGCTT
      TTCACATTTC GGACCCCACG GATTACTCAC TCGATTGAGT GTAATTAACG CAACGCGAGT GACGGGCGAA
3781  TCCAGTCGGG AAACCTGTCG TGCCAGCTGC ATTAATGAAT CGGCCAACGC GCGGGGAGAG GCGGTTTGCG
      AGGTCAGCCC TTTGGACAGC ACGGTCGACG TAATTACTTA GCCGGTTGCG CGCCCCTCTC CGCCAAACGC
3851  TATTGGGCGC TCTTCCGCTT CCTCGCTCAC TGACTCGCTG CGCTCGGTCG TTCGGCTGCG GCGAGCGGTA
      ATAACCCGCG AGAAGGCGAA GGAGCGAGTG ACTGAGCGAC GCGAGCCAGC AAGCCGACGC CGCTCGCCAT
3921  TCAGCTCACT CAAAGGCGGT AATACGGTTA TCCACAGAAT CAGGGGATAA CGCAGGAAAG AACATGTGAG
      AGTCGAGTGA GTTTCCGCCA TTATGCCAAT AGGTGTCTTA GTCCCCTATT GCGTCCTTTC TTGTACACTC
3991  CAAAAGGCCA GCAAAAGGCC AGGAACCGTA AAAAGGCCGC GTTGCTGGCG TTTTTCCATA GGCTCCGCCC
      GTTTTCCGGT CGTTTTCCGG TCCTTGGCAT TTTTCCGGCG CAACGACCGC AAAAAGGTAT CCGAGGCGGG
4061  CCCTGACGAG CATCACAAAA ATCGACGCTC AAGTCAGAGG TGGCGAAACC CGACAGGACT ATAAAGATAC
```

GGGACTGCTC GTAGTGTTTT TAGCTGCGAG TTCAGTCTCC ACCGCTTTGG GCTGTCCTGA TATTTCTATG

4131 CAGGCGTTTC CCCCTGGAAG CTCCCTCGTG CGCTCTCCTG TTCCGACCCT GCCGCTTACC GGATACCTGT

GTCCGCAAAG GGGGACCTTC GAGGGAGCAC GCGAGAGGAC AAGGCTGGGA CGGCGAATGG CCTATGGACA

4201 CCGCCTTTCT CCCTTCGGGA AGCGTGGCGC TTTCTCATAG CTCACGCTGT AGGTATCTCA GTTCGGTGTA

GGCGGAAAGA GGGAAGCCCT TCGCACCGCG AAAGAGTATC GAGTGCGACA TCCATAGAGT CAAGCCACAT

4271 GGTCGTTCGC TCCAAGCTGG GCTGTGTGCA CGAACCCCCC GTTCAGCCCG ACCGCTGCGC CTTATCCGGT

CCAGCAAGCG AGGTTCGACC CGACACACGT GCTTGGGGGG CAAGTCGGGC TGGCGACGCG GAATAGGCCA

4341 AACTATCGTC TTGAGTCCAA CCCGGTAAGA CACGACTTAT CGCCACTGGC AGCAGCCACT GGTAACAGGA

TTGATAGCAG AACTCAGGTT GGGCCATTCT GTGCTGAATA GCGGTGACCG TCGTCGGTGA CCATTGTCCT

4411 TTAGCAGAGC GAGGTATGTA GGCGGTGCTA CAGAGTTCTT GAAGTGGTGG CCTAACTACG GCTACACTAG

AATCGTCTCG CTCCATACAT CCGCCACGAT GTCTCAAGAA CTTCACCACC GGATTGATGC CGATGTGATC

4481 AAGGACAGTA TTTGGTATCT GCGCTCTGCT GAAGCCAGTT ACCTTCGGAA AAAGAGTTGG TAGCTCTTGA

TTCCTGTCAT AAACCATAGA CGCGAGACGA CTTCGGTCAA TGGAAGCCTT TTTCTCAACC ATCGAGAACT

4551 TCCGGCAAAC AAACCACCGC TGGTAGCGGT GGTTTTTTTG TTTGCAAGCA GCAGATTACG CGCAGAAAAA

AGGCCGTTTG TTTGGTGGCG ACCATCGCCA CCAAAAAAAC AAACGTTCGT CGTCTAATGC GCGTCTTTTT

4621 AAGGATCTCA AGAAGATCCT TTGATCTTTT CTACGGGGTC TGACGCTCAG TGGAACGAAA ACTCACGTTA

TTCCTAGAGT TCTTCTAGGA AACTAGAAAA GATGCCCCAG ACTGCGAGTC ACCTTGCTTT TGAGTGCAAT

4691 AGGGATTTTG GTCATGAGAT TATCAAAAAG GATCTTCACC TAGATCCTTT TAAATTAAAA ATGAAGTTTT

TCCCTAAAAC CAGTACTCTA ATAGTTTTTC CTAGAAGTGG ATCTAGGAAA ATTTAATTTT TACTTCAAAA

4761 AAATCAATCT AAAGTATATA TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT GAGGCACCTA

TTTAGTTAGA TTTCATATAT ACTCATTTGA ACCAGACTGT CAATGGTTAC GAATTAGTCA CTCCGTGGAT

4831 TCTCAGCGAT CTGTCTATTT CGTTCATCCA TAGTTGCCTG ACTCCCCGTC GTGTAGATAA CTACGATACG

AGAGTCGCTA GACAGATAAA GCAAGTAGGT ATCAACGGAC TGAGGGGCAG CACATCTATT GATGCTATGC

4901 GGAGGGCTTA CCATCTGGCC CCAGTGCTGC AATGATACCG CGAGACCCAC GCTCACCGGC TCCAGATTTA

CCTCCCGAAT GGTAGACCGG GGTCACGACG TTACTATGGC GCTCTGGGTG CGAGTGGCCG AGGTCTAAAT

4971 TCAGCAATAA ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA GTGGTCCTGC AACTTTATCC GCCTCCATCC

AGTCGTTATT TGGTCGGTCG GCCTTCCCGG CTCGCGTCTT CACCAGGACG TTGAAATAGG CGGAGGTAGG

5041 AGTCTATTAA TTGTTGCCGG GAAGCTAGAG TAAGTAGTTC GCCAGTTAAT AGTTTGCGCA ACGTTGTTGC

TCAGATAATT AACAACGGCC CTTCGATCTC ATTCATCAAG CGGTCAATTA TCAAACGCGT TGCAACAACG

5111 CATTGCTACA GGCATCGTGG TGTCACGCTC GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG TTCCCAACGA

GTAACGATGT CCGTAGCACC ACAGTGCGAG CAGCAAACCA TACCGAAGTA AGTCGAGGCC AAGGGTTGCT

5181 TCAAGGCGAG TTACATGATC CCCCATGTTG TGCAAAAAAG CGGTTAGCTC CTTCGGTCCT CCGATCGTTG

AGTTCCGCTC AATGTACTAG GGGGTACAAC ACGTTTTTTC GCCAATCGAG GAAGCCAGGA GGCTAGCAAC

5251 TCAGAAGTAA GTTGGCCGCA GTGTTATCAC TCATGGTTAT GGCAGCACTG CATAATTCTC TTACTGTCAT

AGTCTTCATT CAACCGGCGT CACAATAGTG AGTACCAATA CCGTCGTGAC GTATTAAGAG AATGACAGTA

5321 GCCATCCGTA AGATGCTTTT CTGTGACTGG TGAGTACTCA ACCAAGTCAT TCTGAGAATA GTGTATGCGG

CGGTAGGCAT TCTACGAAAA GACACTGACC ACTCATGAGT TGGTTCAGTA AGACTCTTAT CACATACGCC

5391 CGACCGAGTT GCTCTTGCCC GGCGTCAATA CGGGATAATA CCGCGCCACA TAGCAGAACT TTAAAAGTGC

GCTGGCTCAA CGAGAACGGG CCGCAGTTAT GCCCTATTAT GGCGCGGTGT ATCGTCTTGA AATTTTCACG

```
5461   TCATCATTGG AAAACGTTCT TCGGGGCGAA AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTCGAT
       AGTAGTAACC TTTTGCAAGA AGCCCCGCTT TTGAGAGTTC CTAGAATGGC GACAACTCTA GGTCAAGCTA
5531   GTAACCCACT CGTGCACCCA ACTGATCTTC AGCATCTTTT ACTTTCACCA GCGTTTCTGG GTGAGCAAAA
       CATTGGGTGA GCACGTGGGT TGACTAGAAG TCGTAGAAAA TGAAAGTGGT CGCAAAGACC CACTCGTTTT
5601   ACAGGAAGGC AAAATGCCGC AAAAAAGGGA ATAAGGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC
       TGTCCTTCCG TTTTACGGCG TTTTTTCCCT TATTCCCGCT GTGCCTTTAC AACTTATGAG TATGAGAAGG
5671   TTTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG AATGTATTTA
       AAAAAGTTAT AATAACTTCG TAAATAGTCC CAATAACAGA GTACTCGCCT ATGTATAAAC TTACATAAAT
5741   GAAAAATAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA AAAGTGCCAC CTGACGTCTA AGAAACCATT
       CTTTTTATTT GTTTATCCCC AAGGCGCGTG TAAAGGGGCT TTTCACGGTG GACTGCAGAT TCTTTGGTAA
5811   ATTATCATGA CATTAACCTA TAAAAATAGG CGTATCACGA GGCCCTTTCG TCTCGCGCGT TTCGGTGATG
       TAATAGTACT GTAATTGGAT ATTTTTATCC GCATAGTGCT CCGGGAAAGC AGAGCGCGCA AAGCCACTAC
5881   ACGGTGAAAA CCTCTGACAC ATGCAGCTCC CGGAGACGGT CACAGCTTGT CTGTAAGCGG ATGCCGGGAG
       TGCCACTTTT GGAGACTGTG TACGTCGAGG GCCTCTGCCA GTGTCGAACA GACATTCGCC TACGGCCCTC
5951   CAGACAAGCC CGTCAGGGCG CGTCAGCGGG TGTTGGCGGG TGTCGGGGCT GGCTTAACTA TGCGGCATCA
       GTCTGTTCGG GCAGTCCCGC GCAGTCGCCC ACAACCGCCC ACAGCCCCGA CCGAATTGAT ACGCCGTAGT
6021   GAGCAGATTG TACTGAGAGT GCACCATATG CGGTGTGAAA TACCGCACAG ATGCGTAAGG AGAAAATACC
       CTCGTCTAAC ATGACTCTCA CGTGGTATAC GCCACACTTT ATGGCGTGTC TACGCATTCC TCTTTTATGG
6091   GCATCAGGCG CCATTCGCCA TTCAGGCTGC GCAACTGTTG GGAAGGGCGA TCGGTGCGGG CCTCTTCGCT
       CGTAGTCCGC GGTAAGCGGT AAGTCCGACG CGTTGACAAC CCTTCCCGCT AGCCACGCCC GGAGAAGCGA
6161   ATTACGCCAG CTGGCGAAAG GGGGATGTGC TGCAAGGCGA TTAAGTTGGG TAACGCCAGG GTTTTCCCAG
       TAATGCGGTC GACCGCTTTC CCCCTACACG ACGTTCCGCT AATTCAACCC ATTGCGGTCC CAAAAGGGTC
6231   TCACGACGTT GTAAAACGAC GGCGCAAGGA ATGGTGCATG CAAGGAGATG GCGCCCAACA GTCCCCCGGC
       AGTGCTGCAA CATTTTGCTG CCGCGTTCCT TACCACGTAC GTTCCTCTAC CGCGGGTTGT CAGGGGGCCG
6301   CACGGGGCCT GCCACCATAC CCACGCCGAA ACAAGCGCTC ATGAGCCCGA AGTGGCGAGC CCGATCTTCC
       GTGCCCCGGA CGGTGGTATG GGTGCGGCTT TGTTCGCGAG TACTCGGGCT TCACCGCTCG GGCTAGAAGG
6371   CCATCGGTGA TGTCGGCGAT ATAGGCGCCA GCAACCGCAC CTGTGGCGCC GGTGATGCCG GCCACGATGC
       GGTAGCCACT ACAGCCGCTA TATCCGCGGT CGTTGGCGTG GACACCGCGG CCACTACGGC CGGTGCTACG
6441   GTCCGGCGTA GAGGCGATTA GTCCAATTTG TTAAAGACAG GATATCAGTG GTCCAGGCTC TAGTTTTGAC
       CAGGCCGCAT CTCCGCTAAT CAGGTTAAAC AATTTCTGTC CTATAGTCAC CAGGTCCGAG ATCAAAACTG
6511   TCAACAATAT CACCAGCTGA AGCCTATAGA GTACGAGCCA TAGATAAAAT AAAAGATTTT ATTTAGTCTC
       AGTTGTTATA GTGGTCGACT TCGGATATCT CATGCTCGGT ATCTATTTTA TTTTCTAAAA TAAATCAGAG
6581   CAGAAAAAGG GGGGAA
       GTCTTTTTCC CCCCTT
```